cDNA Cloning Reveals the Molecular Structure of A Sperm Surface Protein, PH-20, Involved in Sperm–Egg Adhesion And the Wide Distribution of Its Gene among Mammals

William F. Lathrop, Ellen P. Carmichael, Diana G. Myles, and Paul Primakoff

Department of Physiology, University of Connecticut Health Center, Farmington, Connecticut 06032

Abstract. Sperm binding to the egg zona pellucida in mammals is a cell–cell adhesion process that is generally species specific. The guinea pig sperm protein PH-20 has a required function in sperm adhesion to the zona pellucida of guinea pig eggs. PH-20 is located on both the sperm plasma membrane and acrosomal membrane. We report here the isolation and sequence of a full-length cDNA for PH-20 (available from EMBL/GenBank/DDBJ under accession number X56332). The derived amino acid sequence shows a mature protein of 468 amino acids containing six N-linked glycosylation sites and twelve cysteines, eight of which are tightly clustered near the COOH terminus. The sequence indicates PH-20 is a novel protein with no relationship to the mouse sperm adhesion protein galactosyl transferase and no significant homology with other known proteins. The two PH-20 populations, plasma membrane and acrosomal membrane, could arise because one form of PH-20 is encoded and differentially targeted at different spermatogenic stages. Alternatively, two different forms of PH-20 could be encoded. Our evidence thus far reveals only one sequence coding for PH-20: Southern blots of guinea pig genomic DNA indicated there is a single PH-20 gene, Northern blots showed a single size PH-20 message (~2.2 kb), and no sequence variants were found among the sequenced cDNA clones. Cross-species Southern blots reveal the presence of a homologue of the PH-20 gene in mouse, rat, hamster, rabbit, bovine, monkey, and human genomic DNA, showing the PH-20 gene is conserved among mammals. Since genes for zona glycoproteins are also conserved among mammals, the general features of sperm and zona proteins involved in mammalian sperm–egg adhesion may have been evolutionarily maintained. Species specificity may result from limited changes in these molecules, either in their binding domains or in other regions that affect the ability of the binding domains to interact.

ADHESION of sperm to the zona pellucida of the egg is an essential, early step of fertilization in mammals. In previous in vitro studies using guinea pig gametes, we found that both acrosome-intact and acrosome-reacted sperm can initiate adhesion to the egg zona (Myles et al., 1987).

The guinea pig sperm protein PH-20, located on both the plasma membrane and acrosomal membrane, has a required function in acrosome-reacted sperm adhesion to the zona (Primakoff et al., 1985; Myles et al., 1987). Considerable evidence suggests that PH-20 is a sperm adhesion protein and has properties of cell adhesion proteins found on other cell types. An mAb to the PH-20 protein (PH-20 mAb) as well as polyclonal antisera to purified PH-20 inhibit acrosome-reacted sperm adhesion to the zona 80–94%, whereas a control mAb to a different PH-20 epitope and preimmune sera do not inhibit adhesion (Primakoff et al., 1985; Primakoff et al., 1988*b*). Immunization of female guinea pigs with purified PH-20 protein completely blocks in vivo fertility (Primakoff et al., 1988*b*). The PH-20 protein is anchored in the membrane by covalent linkage to glycosyl phosphatidylinositol (GPI)[1] (Phelps et al., 1988) as are various other cell adhesion proteins including N-CAM, F3 and F11 on neurons, Mel-14 on lymphocytes, LFA-3 on antigen presenting cells, and contact site A on Dictyostelium (He et al., 1986; Hemperly et al., 1986; Seed, 1987; Dustin et al., 1987; Low and Saltiel, 1988; Gennarini et al., 1989; Brümmendorf et al., 1989; Camerini et al., 1989). Phosphatidylinositol-specific phospholipase C (PI-PLC) can generally cleave some fraction of a GPI-anchored protein from the surface of cells (Low and Saltiel, 1988). PI-PLC removes ~50% of ^{125}I surface-labeled PH-20 from live acrosome-reacted sperm (Phelps et al., 1988), though no other ^{125}I surface-labeled protein is released (Phelps et al., 1988; our unpublished results). After live acrosome-reacted sperm are treated with PI-PLC and half the PH-20 is released, the sperms' ability to bind to the zona is inhibited 55–60% (Myles, D. G.,

1. *Abbreviations used in this paper*: GPI, glycosyl phosphatidylinositol; PCR, polymerase chain reaction.

and P. Primakoff, unpublished results). Like the adhesion proteins GMP-140 (on endothelial cells and platelets) and CD11b/CD18 and CD11c/CD18 (on monocytes and neutrophils), each of whose surface expression is increased by Ca^{2+}-regulated exocytosis, the surface density of PH-20 is increased about threefold by the acrosome reaction. Thus, the surface concentration of GMP-140, CD11b/CD18, CD11c/CD18, and PH-20 are rapidly upregulated at the time of cell–cell adhesion (McEver and Martin, 1984; Hsu-Lin et al., 1984; Springer et al., 1984; Todd et al., 1984; Cowan et al., 1986; Miller et al., 1987; Kishimoto et al., 1989; Hottori et al., 1989; Geng et al., 1990). All these studies support or are at least consistent with the idea that PH-20 is a sperm adhesion protein that binds to the zona. However, direct evidence demonstrating this has not yet been obtained.

In other mammalian species, various sperm proteins have been proposed to have a role in sperm adhesion to the zona. In the mouse, Shur and his colleagues have shown that a sperm surface galactosyl transferase is an adhesion protein that functions in acrosome-intact sperm binding to the zona (Shur and Hall, 1982; Lopez et al., 1985; Shur and Neely, 1988; Shur, 1989*a*). Bleil and Wassarman (1980) showed that the ligand for acrosome-intact mouse sperm binding is the zona glycoprotein ZP-3. They have reported the presence on the mouse sperm surface of a 56 kD protein which has affinity for ZP-3 (and not ZP-2), is apparently not galactosyl transferase, and is a candidate sperm adhesion protein (Bleil and Wassarman, 1990). Leyton and Saling (1989) have presented evidence that a 95-kD mouse sperm protein that contains phosphotyrosine is a sperm receptor for the zona. These various reports may mean that acrosome-intact mouse sperm have several surface proteins that act in concert as adhesion proteins, but so far there is no evidence to show this is actually the case. On rat sperm, there is a galactose receptor (RTG-r), related to the hepatic asialoglycoprotein receptor, which could function through its lectin properties in sperm binding to zona oligosaccharides (Abdullah and Kierszenbaum, 1989). Another rat sperm surface antigen, 2B1, has been proposed to play a role in rat sperm–zona adhesion (Shalgi et al., 1990; Jones et al., 1990). A boar sperm plasma membrane protein (AP_z), distinct from galactosyl transferase, and a rabbit sperm protein (RSA) have also been reported to have a role in sperm–zona adhesion (Peterson and Hunt, 1989; O'Rand et al., 1988).

The diversity of species studied and multiplicity of candidate adhesion proteins has produced uncertainty about the actual mechanism of sperm-zona adhesion (for review see Saling, 1989; Shur, 1989*b*). Two unanswered questions contribute to this uncertainty. First, do different mammalian species have the same sperm proteins functioning in sperm–zona adhesion? Mammalian sperm-zona adhesion is generally a species-specific process. For instance, both acrosome-intact and acrosome-reacted guinea pig sperm will adhere to the zona of guinea pig eggs but not to the zona of other rodent (mouse and hamster) eggs (Schmell and Gulyas, 1980; Yanagimachi, 1981; Myles et al., 1987). Thus, it is conceivable that in some cases different mammals use different sperm adhesion proteins for the zona. Second, are any of the candidate sperm adhesion proteins from different species actually the same, i.e., coded by evolutionarily closely related, homologous genes? It seems possible, on the basis of molecular mass, that guinea pig PH-20 (~64 kD) could be the guinea pig homologue of mouse galactosyl transferase (60 kD), the mouse 56 kD protein that binds ZP-3, rat RTG-r (54 kD), or boar APz (55 kD), and is less likely related to mouse 95 kD rat 2B1 (40 kD), and rabbit RSA (14–17 kD).

To provide a definitive answer to the question of PH-20's relationship to proposed adhesion proteins on sperm from other species as well as to adhesion proteins of other cell types, we have cloned and sequenced a cDNA for guinea pig PH-20. The sequence reveals that PH-20 is a novel protein with no significant homology with other known proteins. Cross-species Southern blots show there is a homologue of the PH-20 gene in a wide range of mammals suggesting that the general features of PH-20 are conserved among mammals. Species specificity of adhesion may arise through limited changes in adhesion molecules either in their binding domains or in other regions that influence the ability of the binding domains to interact.

Materials and Methods

Tryptic Peptide Sequence Determination

PH-20 protein was purified on an mAb affinity column as described (Primakoff et al., 1988*a*). The purified PH-20 was digested with trypsin, the resulting peptides were separated by HPLC, and selected peptides sequenced (Stone et al., 1989).

Library Construction and Screening

A population of testicular cells, enriched for spermatogenic cells on a Percoll gradient (Phelps et al., 1988), was used for the isolation of spermatogenic cell total RNA using a modification of the procedure described by Ausubel et al. (1987). The pelleted cells were lysed with detergent in the presence of vanadyl–ribonucleoside complexes (VRCs) in 0.5–1.0 ml of solution containing 10 mM Tris, pH 8.6, 0.5% NP-40, 0.14 M NaCl, 1.5 mM $MgCl_2$, and 10 mM VRC. After pelleting cellular debris, 0.5 vol of 2× Proteinase K buffer (2× = 0.2 M Tris, pH 7.5, 25 mM EDTA, pH 8.0, 0.3 M NaCl, and 2.0% SDS) and 200 μg/ml Proteinase K (Sigma Chemical Co., St. Louis, MO) were added to the supernatant. PolyA$^+$ RNA was purified from the total RNA by oligo-dT cellulose chromatography essentially as described in Maniatis et al. (1982). cDNA was synthesized by a modification of the method of Gubler and Hoffman (1983), ligated with lambda gt11 arms, and packaged into lambda coat proteins, utilizing kits and protocols from Amersham Corp. (Arlington Heights, IL).

The unamplified library was plated at 20,000 plaques/150-mm plate for screening. A single nitrocellulose filter from each plate was immunoblotted with rabbit anti-PH-20 polyclonal antiserum, raised against affinity-purified PH-20 protein (Primakoff et al., 1988*a*), and diluted 1/500 in TBST (10 mM Tris, pH 8.0, 0.15 M NaCl, 0.05% Tween-20) containing 2 mg/ml *E. coli* protein. The *E. coli* protein was prepared by pelleting an overnight culture of Y1090 cells, resuspending the cells in a minimal volume of TBST and freezing in liquid nitrogen. The thawed cells were sonicated and the protein concentration determined using the BCA reagent (Pierce Chemical Co., Rockford, IL). Positive plaques were detected with an anti-rabbit IgG alkaline phosphatase–conjugated second antibody (Promega Biotec, Madison, WI). Size of the fusion protein made by plaque-purified positive clones was determined by analyzing *E. coli* extracts containing the fusion protein on SDS-PAGE (Mierendorf et al., 1987). Inserts from positive clones were subcloned into pUC19.

Polymerase Chain Reaction Generation of the 5′ Portion of the cDNA

The 5′ portion of the PH-20 cDNA was cloned utilizing anchored polymerase chain reaction (PCR) following the protocol of Frohman et al. (1988). PolyA$^+$ RNA from spermatogenic cells (2 μg in 10 μl dH_2O) was heated to 65°C for 3 min and then reverse transcribed as described (Frohman et al., 1988) by adding 4 μl of 10× RTC buffer (1× buffer is 50 mM Tris, pH 8.3, 50 mM KCl, 4 mM dithiothreitol, 10 mM $MgCl_2$), and 4 μl 10 mM stock of each dNTP (1 mM final), 2 μl of 80 mM sodium pyrophosphate

(4 mM final), 1 μl (40 U) of RNasin (Promega Biotec), 40 pmol PH-20 specific primer (PH-20-RT), 18 U AMV reverse transcriptase (Life Sciences, St. Petersburg, FL), and 40 μCi 32P-dCTP in 40 μl total volume. After 1 h of incubation at 42°C, an additional 1 μl of reverse transcriptase was added and incubation continued for a second hour. The PH-20-RT primer was a 17 nucleotide (nt) oligomer (nt 1,242–1,258, see Fig. 1 *b*), ~250 bases downstream from the 5′ end of the insert gpPH-20-1 (see Fig. 1 *a*). The single strand cDNA was separated from excess PH-20-RT by column chromatography, tailed with polyA, and diluted to 1.0 ml (Frohman et al., 1988). Second strand synthesis and PCR amplification were performed with a GeneAmp kit (Perkin Elmer Corp., Norwalk, CT) in a 100-μl reaction containing 10 μl of the reverse transcription product, 20 pmol (dT)17-adapter, 50 pmol adapter, and 50 pmol PH-20-AMP primer, as described by Frohman et al. (1988). The PH-20-AMP primer was a 17-nt oligomer (nt 1,202–1,218, Fig. 1 *b*) located upstream from the PH-20-RT primer. The PCR product was purified from unincorporated primers and free nucleotides by spin column chromatography (columns from Boehringer-Mannheim Biochemicals, Indianapolis, IN). It was subsequently digested with Hgi AI and Sal I, gel purified, and ligated into pBluescript digested with Pst I and Sal I.

DNA Sequencing

DNA sequence was determined using the dideoxynucleotide chain termination method and Sequenase (United States Biochemical Corp., Cleveland, OH). Plasmid DNA was isolated by alkaline lysis and purified by CsCl gradient centrifugation (Maniatis et al., 1982). The double-stranded template was sequenced using site specific primers (17 nt long).

Western Blot of Eluted Antibody

A plate showing confluent lysis by the lambda clone containing the insert gpPH-20-1 (see Fig. 1 *a*) was overlaid with a nitrocellulose filter soaked in 10 mM isopropylthiogalactoside (IPTG) and protein allowed to bind to the filter. Rabbit anti-PH-20 polyclonal antiserum was bound to the beta-galactosidase-PH-20 fusion protein and then eluted with 0.5 M acetic acid, 0.5 M NaCl, pH 2.5, neutralized, concentrated, and used for Western blotting.

Northern Blot

PolyA$^+$ RNA from spermatogenic cells was separated by electrophoresis on a 0.8% formaldehyde gel and transferred to nitrocellulose (Schleicher & Schuell, Inc., Keene, NH) (Maniatis et al., 1982). The nitrocellulose was prehybridized in 50% formamide, 4×SSC (1×SSC = 150 mM NaCl, 15 mM sodium citrate), 50 mM $NaPO_4$, pH 6.5, 0.1% SDS, 100 μg/ml salmon sperm DNA, and 1× Denhardt's solution (0.02% polyvinylpyrrolidine, 0.02% BSA, 0.02% Ficoll) for 2 h at 65°C. The hybridization buffer (prehybridization buffer with 20% dextran sulfate) plus 1 × 10^6 cpm/ml probe was added and incubated overnight at 65°C. Probe was prepared by the random hexamer method (using a kit and protocol from Bio-Rad laboratories, Richmond, CA) and insert gpPH-20-1 as template. The nitrocellulose was washed in 2×SSC + 0.1% SDS 4 × 5 min at room temperature then 4 × 15 min in 0.1×SSC + 0.1% SDS at 55°C. The nitrocellulose was dried and exposed to film at −70°C with an intensifying screen.

Southern Blot Analysis

Genomic DNA was isolated from guinea pig, rat, rabbit, mouse, and hamster spleens by detergent lysis-Proteinase K digestion (Ausubel et al., 1987). Some DNA samples were kindly provided by other investigators at the University of Connecticut Health Center: the human DNA was a gift from Dr. Petros Tsipouras; monkey DNA, isolated from Vero cells, was a gift from Dr. Sandra Weller; and chicken DNA was a gift from Ms. Hyun-Duck Nah. DNA from salmon sperm and bovine thymus were purchased from Sigma Chemical Co. and reconstituted at 1 mg/ml in TE (10 mM Tris, pH 8.0, 1 mM EDTA, pH 8.0). All species' DNA (10 μg) were cut with restriction enzymes and separated on a 1% agarose gel. The Southern transfer was carried out by capillary transfer (Maniatis et al., 1982) onto Magnagraph nylon membrane (Micron Separations, Inc, Westboro, MA). The membranes were prehybridized in a solution consisting of 6×SSC, 1× Denhardt's, 250 mg/ml salmon sperm DNA, 1% SDS, and 50 mM $NaPO_4$, pH 7.4, for 1–2 h at 65°C. The membranes were hybridized overnight at 55°C in prehybridization buffer plus 2 × 10^6 cpm/ml probe. Probes were prepared by the random hexamer method using (unless stated otherwise) inserts gpPH-20-1 and gpPH-20-3 as templates in separate reactions and then pooling the two reaction mixes. The blot was washed 3 × 5 min in 2× SSC + 0.1% SDS at room temperature, 2 × 30 min in 2× SSC + 0.1% SDS at 50°C, and 2 × 30 min in 1× SSC + 0.1% SDS at 60°C. The blot was wrapped in plastic wrap and exposed to film with an intensifying screen at −70°C.

Computer Searches for Sequence Homologies

The Protein Identification Resource data base, Release 21 of the National Biomedical Research Foundation (NBRF), and the EMBL database were searched for sequences related to PH-20 using the FASTP and TFASTA programs. In addition, PH-20's derived amino acid sequence was compared to the amino acid sequences of the alpha and beta chains of CD11c/CD18 using programs on the NBRF package.

Results

Isolation and Sequencing of cDNAs

Using size-selected cDNA (0.5–7 kb) made from spermatogenic cell polyA$^+$ RNA, we constructed a lambda gt11 library containing 9 × 10^5 independent recombinants. Screening with anti-PH-20 polyclonal antibody detected six positive recombinants that were plaque purified. Insert size of these clones ranged from 0.09 to 2.2 kb. The size of the beta-galactosidase-insert fusion proteins, produced by these clones, varied from 118–157 kD with two of the clones making no detectable polypeptide fused with beta-galactosidase (data not shown). All six inserts were sequenced at least partially. Two of the inserts were confirmed to code for the PH-20 protein by locating the sequences of two PH-20 tryptic peptides (Table I, peptides #2 and #3) in their derived amino acid sequence (Fig. 1 *b*). Both of these inserts (gpPH-20-1, nt 1,016–2,152 and gpPH-20-2, nt 1,010–2,125, Fig. 1, *a* and *b*) contained a long (~925 nt) open reading frame, a stop codon, a 3′ untranslated region, and a polyA tail. Thus, these two inserts were concluded to represent the 3′ end of a cDNA for PH-20. The other four antibody-positive lambda clones were unrelated to PH-20.

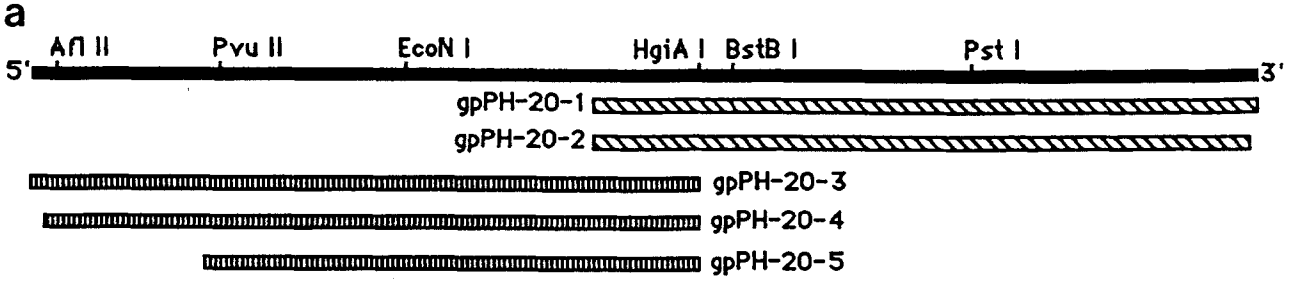
To obtain the 5′ portion of PH-20 cDNA we utilized an anchored PCR method (Frohmann et al., 1988). PolyA$^+$ RNA from spermatogenic cells was reverse transcribed using a primer (PH-20-RT) whose sequence was derived from the 5′ end of the insert gpPH-20-1. The single strand cDNA was tailed and amplified using primers complementary to the tail and another primer (PH-20-AMP) from the 5′ end of the insert gpPH-20-1 (see Materials and Methods). The major PCR product was 1.2 kb, and Southern blot analysis

Table I. Tryptic Peptides Isolated from PH-20 Protein

Peptide #	Peptide sequence
1	kAFMLETLK
2	rSLVSCIGLENYMK
3	kFVVHGKPSLEDLQEFSK
4	rAPPLIPNVPLXXVXNAPXEFCI

PH-20 protein was purified as described (Primakoff et al., 1988*a*), then digested with trypsin, and the resulting peptides were separated by HPLC and selected peptides sequenced (Stone et al., 1989). The *r* or *k* (the amino acids after which trypsin cleaves) shown preceding the peptide sequence are the residues determined for that position in the derived amino acid sequence. *X* denotes positions at which no residue could be determined in the peptide sequencing.

a



b

```
GCTTTACTGTGAGGTTGCTTGTACATTGATTTTCCAGTTCTCTTAAGAATCTGTGGCTTGATGTAGCTCACACGAATCCAGGAGGATTTTGGTTTCTTA   99
ATTTTGATGACTGCGTACATGATTAGTAGTACATCGTAAAGTCTCTTCCAACAAGTTACAGATGGTGCAACATTCAAAACATTCCTGAAATACAAAACA  198
AGAAGAATATTTTAATGTAACAGAGTTGTTTACCTCTTTATCCACCAAAGTGACCTCACTGTACTACGCTTCTTTTGGGCTCATATTGTGCAACAAATA  297
TTGGAAAAAACAGTGTATAAGAAGAAAAAGTATTTTTCACAGCTGTTACTCTTTCTAATGGGAGCATTCACTTTTAAACACAGCTTTTTTGGGAGTTTT  396
                                                  M  G  A  F  T  F  K  H  S  F  F  G  S  F   -22

GTTGAGTGCAGTGGAGTATTGCAGACAGTGTTTATTTTCCTTCTAATTCCATGTTGCCTGGCTGATAAAAGGGCACCACCACTCATCCCAAATGTGCCT  495
 V  E  C  S  G  V  L  Q  T  V  F  I  F  L  L  I  P  C  C  L  A  D  K  R  A  P  P  L  I  P  N  V  P   12
                                                                  1

TTGCTCTGGGTCTGGAATGCCCCAACTGAATTTTGTATAGGAGGAACCAATCAACCACTAGATATGAGCTTTTTCTCTATAGTAGGAACTCCCAGGAAA  594
 L  L  W  V  W  N  A  P  T  E  F  C  I  G  G  T  N  Q  P  L  D  M  S  F  F  S  I  V  G  T  P  R  K   45
                                    X

AATATCACAGGGCAAAGTATTACACTATATTATGTTGATAGACTTGGCTACTATCCTTACATAGATCCTCACACAGGTGCGATTGTGCATGGAGGACTC  693
 N  I  T  G  Q  S  I  T  L  Y  Y  V  D  R  L  G  Y  Y  P  Y  I  D  P  H  T  G  A  I  V  H  G  G  L   78
 o

CCCCAGTTGATGAATTTACAACAGCATTTGAGAAAATCAAGGCAAGACATTTTATTTTACATGCCCACAGACAGCGTGGGCTTGGCTGTCATTGACTGG  792
 P  Q  L  M  N  L  Q  Q  H  L  R  K  S  R  Q  D  I  L  F  Y  M  P  T  D  S  V  G  L  A  V  I  D  W  111

GAAGAGTGGAGGCCCACCTGGACAAGAAACTGGAGACCTAAGGATATTTACAGGAATAAATCTATTGAGTTGGTTAAGAGCCAGCATCCACAGTATAAT  891
 E  E  W  R  P  T  W  T  R  N  W  R  P  K  D  I  Y  R  N  K  S  I  E  L  V  K  S  Q  H  P  Q  Y  N  144
                                                  o                                             o

CACTCATATGCTGTTGCCGTAGCCAAAAGAGACTTTGAAAGGACAGGGAAGGCTTTCATGCTAGAAACTTTAAAACTGGGAAAATCACTTAGGCCAAGT  990
 H  S  Y  A  V  A  V  A  K  R  D  F  E  R  T  G  K  A  F  M  L  E  T  L  K  L  G  K  S  L  R  P  S  177

AGCTTATGGGGTTATTATCTTTTTCCTGATTGCTACAACACTCATTTCACTAAACCCAATTATGATGGGCATTGCCCTCCTATAGAACTGCAAAGAAAT 1089
 S  L  W  G  Y  Y  L  F  P  D  C  Y  N  T  H  F  T  K  P  N  Y  D  G  H  C  P  P  I  E  L  Q  R  N  210
                              X                                                 X

AACGATCTCCAATGGTTGTGGAACGACAGCACTGCCCTTTACCCATCTGTTTATTTGACCAGTCGAGTAAGATCATCTCAAAATGGTGCACTTTATGTT 1188
 N  D  L  Q  W  L  W  N  D  S  T  A  L  Y  P  S  V  Y  L  T  S  R  V  R  S  S  Q  N  G  A  L  Y  V  243
                        o

CGTAATCGTGTACACGAGTCCATTAGGGTTTCGAAACTCATGGATGACAAAAACCCACTTCCGATTTATGTGTATATCCGCCTCGTTTTTACCGATCAA 1287
 R  N  R  V  H  E  S  I  R  V  S  K  L  M  D  D  K  N  P  L  P  I  Y  V  Y  I  R  L  V  F  T  D  Q  276

ACTACTACATTCCTTGAACTGGATGATCTTGTGCATTCAGTTGGCGAAATTGTTCCTCTAGGTGTCTCTGGAATAATAATATGGGGAAGTCTTAGTTTA 1386
 T  T  T  F  L  E  L  D  D  L  V  H  S  V  G  E  I  V  P  L  G  V  S  G  I  I  I  W  G  S  L  S  L  309

ACACGAAGTTTGGTTTCTTGTATAGGACTAGAAAATTACATGAAGGGTACACTCCTGCCTTACTTAATCAATGTCACCCTAGCAGCCAAAATGTGTGGC 1485
 T  R  S  L  V  S  C  I  G  L  E  N  Y  M  K  G  T  L  L  P  Y  L  I  N  V  T  L  A  A  K  M  C  G  342
                  X                                                   o                    X

CAAGTGCTTTGTAAGAATCAAGGAATTTGCACAAGGAAAGACTGGAACACAAACACCTATCTTCACCTAAACGCAACAAATTTTGACATTGAACTTCAG 1584
 Q  V  L  C  K  N  Q  G  I  C  T  R  K  D  W  N  T  N  T  Y  L  H  L  N  A  T  N  F  D  I  E  L  Q  375
          X              X                                   o

CAAAATGGGAAGTTTGTAGTACATGGAAAACCATCACTTGAAGACCTGCAGGAATTTTCCAAAAATTTTCATTGCTCCTGTTATACCAATGTGGCTTGT 1683
 Q  N  G  K  F  V  V  H  G  K  P  S  L  E  D  L  Q  E  F  S  K  N  F  H  C  S  C  Y  T  N  V  A  C  408
                                                                  X     X                 X

AAGGACAGACTTGATGTACATAATGTCCGTTCTGTTAATGTGTGTACTGCCAATAATATTTGTATAGATGCTGTTTTAAATTTTCCATCCCTGGATGAT 1782
 K  D  R  L  D  V  H  N  V  R  S  V  N  V  C  T  A  N  N  I  C  I  D  A  V  L  N  F  P  S  L  D  D  441
                                             X              X

GATGATGAGCCTCCCATCACAGACGACACGTCTCAAAATCAAGACAGCATCTCCGACATCACATCATCTGCTCCACCGTCTTCACATATTCTTCCAAAG 1881
 D  D  E  P  P  I  T  D  D  T  S  Q  N  Q  D  S  I  S  D  I  T  S  S  A  P  P  S  S  H  I  L  P  K  474

GATCTCAGTTGGTGCCTCTTCCTTCTCTCTATCTTTTCACAGCACTGGAAATATTTACTATAGGCTCATGGCAACTGAAAAGTACCAAATTCAATATCA 1980
 D  L  S  W  C  L  F  L  L  S  I  F  S  Q  H  W  K  Y  L  L  Z                                       495

TAAAAATTCTATAATCAAAATCCTTTGAATTTTTAAAGCAAAATACATACTATTCTATCAAAGACACTGTAAAGCCTGTGGTACTTGGAAGATACAGCT 2079
TTCTTTTGAGAAGAGTGAAGATTTGAATAAAACAAAATTACTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA                           2152
                        *******
```

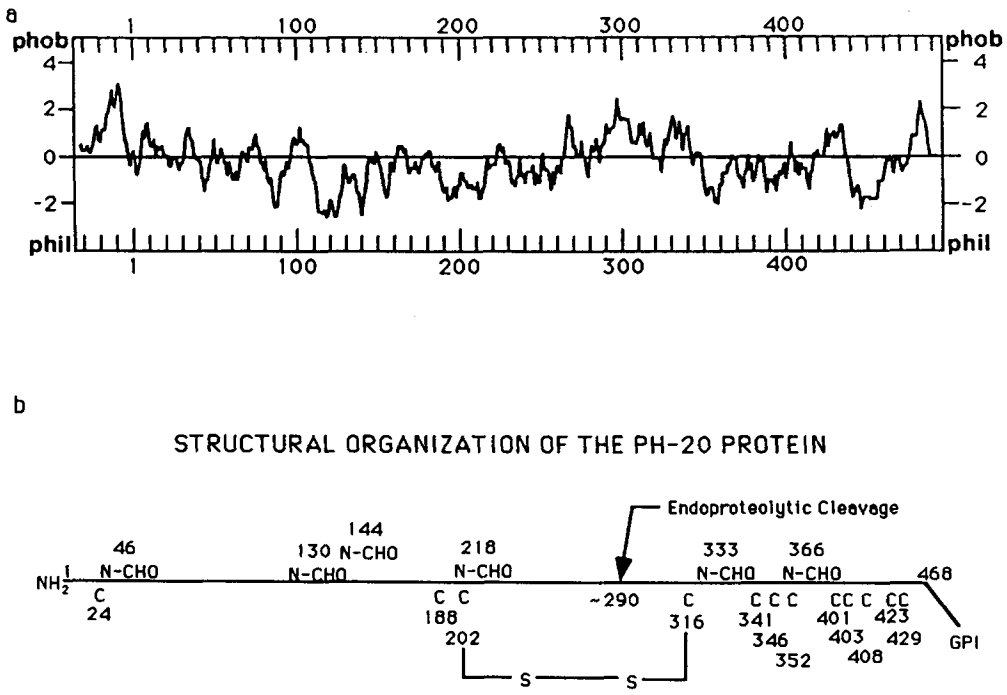


Figure 2. Hydropathy plot and proposed structure of the PH-20 protein. (*a*) Hydropathy plot of the derived amino acid sequence was computed by DNA Strider using the Kyte-Doolittle (Kyte and Doolittle, 1981) algorithm. (*phob*) Hydrophobic; (*phil*) hydrophilic. Amino acid number 1 is the first residue after the proposed signal sequence cleavage. (*b*) A schematic summarizing key structural features of the predicted, mature PH-20 protein. (*N-CHO*) Potential N-linked carbohydrate; (*C*) cysteine; (*GPI*) glycosyl phosphatidylinositol. The schematic shows the predicted attachment of serine 468 to GPI and arbitrarily shows a single disulfide bond between cysteine 202 and cysteine 316 to indicate the linkage of the two endoproteolytic fragments by disulfide bonding.

confirmed that this band hybridized with the labeled insert gpPH-20-1. The major PCR products from three separate reactions were cloned and one insert from each of the three reactions was sequenced (gpPH-20-3, nt 1–1,175, gpPH-20-4, nt 24-1,175, and gpPH-20-5, nt 295-1,175; Fig. 1, *a* and *b*).

The complete cDNA sequence and the derived amino acid sequence (Fig. 1 *b*) were obtained from the five cDNA inserts (Fig. 1 *a*) that were sequenced in their entirety on both strands. The cDNA sequence contains a 354-nt 5′ untranslated region, a 1,590-nt open reading frame, and a 208-nt 3′ untranslated region. The derived amino acid sequence contains all the tryptic peptide sequences obtained from purified PH-20, confirming that the cDNAs are authentic PH-20 clones. As mentioned, the inserts gpPH-20-1 and gpPH-20-2 code for tryptic peptides #2 and #3 (Fig. 1, *a* and *b*; Table I). The inserts gpPH-20-3, gpPH-20-4, and gpPH-20-5 code for tryptic peptides #1 and #4 (Fig. 1, *a* and *b*; Table I).

Structural Characteristics of the PH-20 Protein

The PH-20–derived amino acid sequence has features typical of an integral membrane protein, namely a signal sequence, a sequence for a membrane anchor and sites for N-linked glycosylation. A hydrophobicity plot (Fig. 2 *a*) shows a strongly hydrophobic NH_2-terminal sequence, an internal hydrophobic region (centered around amino acid 295), and a short COOH-terminal hydrophobic region. The NH_2-terminal hydrophobic region has the characteristics of a signal sequence including a hydrophobic core, small uncharged amino acids in the −3 and −1 positions, and an aspartate in the +1 position, given a predicted cleavage between alanine (−1) and aspartate (+1) (Fig. 1 *b*) (von Heijne, 1986). The assignment of aspartate as the NH_2-terminal amino acid of mature PH-20 is a prediction based on analogy to the known sites of cleavage of other signal sequences (von Heijne, 1986; von Heijne, 1987; von Heijne, 1990). Confirmation that cleavage does occur between alanine (−1) and aspartate (+1) has been obtained by NH_2-terminal sequencing of purified PH-20 (our unpublished results). The signal sequence is long (35 residues) compared to the more commonly found ~25 residue signal sequence, though other membrane proteins have similar length signal sequences (e.g., 38 residues for the lymphocyte homing receptor, Mel-14; Siegelman et al., 1989; Laskey et al., 1989).

The COOH-terminal hydrophobic sequence (13 residues; Fig. 1 *b*) is too short to be a transmembrane anchor. This

Figure 1. Partial restriction map, cDNA clones, and sequence for PH-20 cDNA. (*a*) Schematic map showing the location of cDNA clones in respect to the composite full-length sequence. Some restriction enzymes that cut at unique sites are indicated. Inserts gpPH-20-1 (nt 1,016–2,152) and gpPH-20-2 (nt 1,010–2,125) were isolated from a cDNA expression library in lambda gt11 by immunoscreening with an anti-PH-20 antiserum. Inserts gpPH-20-3 (nt 1–1,175), gpPH-20-4 (nt 24–1,175), and gpPH-20-5 (nt 295–1,175) were cloned from three separate PCR amplifications of the 5′ portion of PH-20 mRNA. All inserts were sequenced completely on both strands. (*b*) Nucleotide and derived amino acid sequence of the PH-20 cDNA. The signal sequence corresponds to amino acids -35 to -1. The number 1 denotes the first amino acid of the predicted mature protein. A hydrophobic region potentially signaling GPI anchorage is boxed. Tryptic peptides from purified PH-20 protein are underlined. Possible N-linked glycosylation sites are marked with *o* and cysteine residues in the mature protein with *X*. The stop codon is shown by *Z*. The polyadenylation signal is underlined with *. These sequence data are available from EMBL/GenBank/DDBJ under accession number X56332.

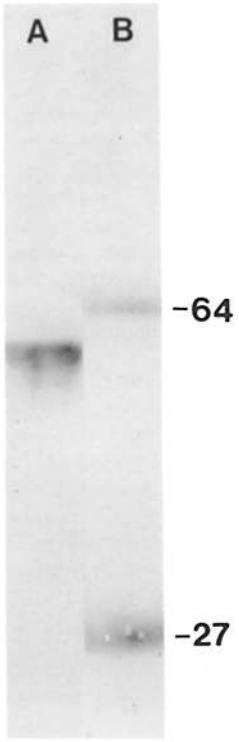


Figure 3. Western blot of purified PH-20 protein. PH-20 protein (200 ng/lane), purified from acrosome-reacted sperm, was analyzed on SDS-PAGE nonreduced (lane *1*) and reduced (lane *2*). The first antibody used to stain the blot was affinity purified from the fusion protein of the lambda gt11 clone with the insert gpPH-20-1 (see Materials and Methods). Molecular masses are shown in kD for the bands in the reduced sample. (lane *1*) The nonreduced sample shows a single band because the proteolytic fragments of PH-20 are disulfide bonded to each other. Nonreduced PH-20 migrates slightly farther on SDS-PAGE than reduced, intact PH-20 (~64 kD) (Primakoff et al., 1988*a*). (lane *2*) The reduced sample shows intact ~64 kD and one of the fragments, ~27 kD. On silver-stained gels, an additional band at 41–48 kD is observed in the reduced sample although it is not detected with this affinity-purified antibody staining this load of purified PH-20 protein.

type of short, hydrophobic COOH-terminal region is characteristic of the COOH-terminal sequences of all sequenced GPI-anchored proteins, suggesting that this short hydrophobic sequence may be the signal for GPI attachment (Ferguson and Williams, 1988; Low, 1989). Based on analogy to the sequences of other GPI-anchored proteins, we predict removal of the COOH-terminal hydrophobic sequence by a potential cleavage between serine 468 and 469. Thus, in mature PH-20, serine 468 is the predicted COOH-terminal amino acid that would be covalently linked to the GPI in the outer leaflet of the bilayer (Ferguson and Williams, 1988; Low, 1989). The mature protein's primary sequence is 468 residues long after the predicted cleavages of the signal sequence (35 residues) and GPI anchor sequence (27 residues) and has a calculated molecular weight of 53,316.

The COOH-terminal sequence for GPI anchoring is consistent with our previous finding that PH-20 on sperm is GPI anchored (Phelps et al., 1988), and suggests that PH-20 has a single extracellular domain and lacks a transmembrane or cytoplasmic domain. Six sites for N-linked glycosylation (Fig. 1 *b*) are fairly evenly spaced along the 468 residue extracellular domain (compare the schematic of proposed structural organization of PH-20, Fig. 2 *b*). Previous work has shown that PH-20 protein purified from cauda epididymal sperm has an M_r ~64 on SDS-PAGE (Primakoff et al., 1988*a*). After purified PH-20 is digested to completion with N-glycanase, it migrates on SDS-PAGE with M_r ~50–52 (our unpublished observations) suggesting that several or all of the N-linked glycosylation sites are used.

Mature PH-20 contains twelve cysteines (Fig. 1 *b*) which suggests the possibility of extensive folding due to intermolecular disulfides. Eight of the cysteines are tightly clustered in the 88 amino acids between residues 341–429, i.e., close to the COOH-terminus (Fig. 2 *b*). Presumably, at least four of the eight clustered cysteines must form disulfides with each other since there are only four other cysteines in the sequence. Thus, one can predict the presence in PH-20 of two or more small loops, held together by disulfides, very close to the lipid bilayer.

Previously we have found that almost all the purified PH-20 protein from acrosome-reacted sperm has been endoproteolytically converted from the intact, M_r ~64, form into two disulfide-linked fragments, migrating on SDS-PAGE with M_r 41–48 and ~27 (Primakoff et al., 1988*a*; Primakoff, P., and D. Myles, unpublished results). PH-20 on the plasma membrane is endoproteolytically cleaved during sperm development in the epididymis and PH-20 on the acrosomal membrane is cleaved after the acrosome reaction; we have speculated that endoproteolysis could activate PH-20 (Phelps et al., 1990; Primakoff et al., 1988*a*). The order of the two proteolytic fragments (41–48 kD and ~27 kD) in intact PH-20 had not been previously determined. To determine if the PH-20 ~27-kD fragment is NH_2-terminal or COOH-terminal, we could use the lambda clone with the insert gp PH-20-1, which codes only for the COOH-terminal half of PH-20. If ~27 kD is NH_2-terminal, it should not be encoded by gpPH-20-1; if ~27 kD is COOH-terminal, it should be encoded by gpPH-20-1. When anti-PH-20 polyclonal antibody was eluted from the fusion protein produced by the lambda clone with the insert gpPH-20-1, the eluted antibody bound strongly to the ~27-kD fragment, showing that ~27 kD is COOH-terminal (Fig. 3). From the approximate size of the two fragments (~44 and ~27 kD), we can make the rough estimate that the endoproteolytic cleavage occurs toward the COOH-terminus at approximately amino acid 290. Cleavage at amino acid 290 would produce a large (~44 kD) fragment (residues 1–290) and a small (~27 kD) fragment (residues 291–468), each of which contain cysteines that could form the disulfide bond(s) that connect them (Fig. 2 *b*).

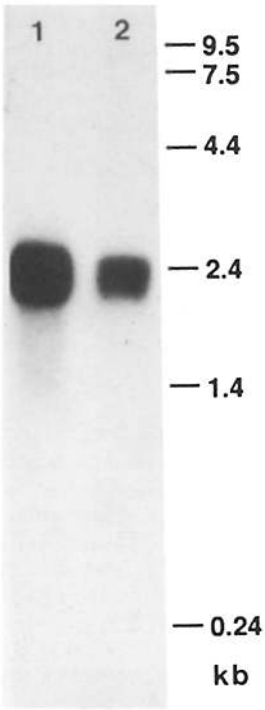


Figure 4. Northern blot of guinea pig spermatogenic cell poly A^+ RNA. Size markers are shown in kb; lane *1* contains 5 μg RNA and lane 2, 1 μg RNA. The probe was made using gpPH-20-1 as a template.

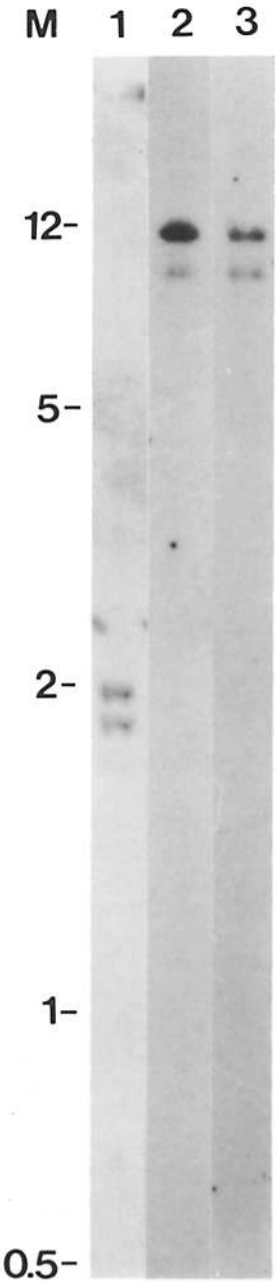


Figure 5. Southern blot of guinea pig genomic DNA. Guinea pig genomic DNA (10 μg/lane) was digested with Sau96 I (lane *1*) and Pst I (lanes *2* and *3*) and analyzed on a 1% agarose gel. Lanes *1* and *2* were hybridized with a mix of labeled gpPH-20-3 and gpPH-20-2 and lane *3* with labeled gpPH-20-2. Size markers (*M*) are in kb.

Sequence Similarity of PH-20 with Other Proteins

Computer searches revealed no significant homology of the PH-20 amino acid sequence with other known sequences. An intriguing short sequence similarity was found between PH-20 and the leukocyte adhesion protein CD11c/CD18 (also called p150/95; Corbi et al., 1987; Kishimoto et al., 1987; Anderson and Springer, 1987). CD11c/CD18 has two noncovalently associated subunits, termed alpha and beta, each of which has a large extracellular domain. One PH-20 sequence, amino acids 210–240, is similar to amino acids 164–194 in the extracellular domain of the alpha subunit. In these 31 residues, there are eight (26%) identical and seven (23%) conserved amino acids. A second PH-20 sequence, amino acids 84–118, is similar to amino acids 124–158 in the extracellular domain of the beta subunit. Among these 35 residues, 11 (31%) are identical and 9 (26%) are conserved.

Evidence that Only One Form of PH-20 Is Encoded

Since in spermatogenic cells and sperm, there are two populations of PH-20, one on the plasma and one on the acrosomal membrane (Cowan et al., 1986; Phelps and Myles, 1987), it seemed possible that two distinct messages or genes coding for PH-20 might be found. However, using polyA$^+$ RNA from spermatogenic cells, only a signal band (~2.2 kb) was detected in a Northern blot (Fig. 4). When guinea pig genomic DNA was digested with the restriction enzymes Sau96 I or Pst I, each of which have a single recognition site in the cDNA sequence, two bands were detected in a Southern blot (Fig. 5). The Sau96 I- and Pst I-digested DNA samples were probed with a mix of labeled gpPH-20-3 (5′ portion of the cDNA) and gpPH-20-2 (3′ portion of the cDNA). The two bands in the Sau96 I digest had about equal intensity (Fig. 5, lane *1*), whereas the shorter hybridizing band in the Pst I digest gave a relatively weak signal (Fig. 5, lane *2*). We reasoned weak hybridization would occur, using the mixed probe, with the 3′ fragment of the gene since the Pst I site is near the 3′ end of the cDNA. Thus, we also used a probe of only labeled gpPH-20-2, the 3′ portion of the cDNA which contains the Pst I site at about the middle of its sequence. With this probe, we found the signal from the shorter hybridizing band was relatively stronger (Fig. 5, lane *3*). The presence of two bands in each digest indicates there is a single gene for PH-20 in guinea pig genomic DNA.

A Homologue of the PH-20 Gene in Other Species

To determine if there is a homologue of the PH-20 gene in the genomic DNA of other species, we did cross-species Southern blots. The blots were probed with a mix of labeled gpPH-20-3 and gpPH-20-2. The Southerns showed a weakly hybridizing band at ~10 kb for chicken DNA and strongly hybridizing bands for mouse, rat, hamster, rabbit, and human DNA (Fig. 6). In addition, hybridization was observed with bovine and monkey DNA (data not shown).

Discussion

Targeting of PH-20 to Two Membranes

Our results provide an initial molecular characterization of the sperm antigen PH-20 which has been implicated as a sperm adhesion protein for the zona pellucida. PH-20 is present in guinea pig sperm on both the posterior head plasma membrane and inner acrosomal membrane. The acrosome reaction results in a threefold increase in PH-20 surface density at the time PH-20 will function in adhesion of acrosome-reacted sperm to the zona (Cowan et al., 1986). One long-term goal in cloning PH-20 cDNA is to understand the mechanism for targeting PH-20 to the plasma and acrosomal membranes in round spermatids, where it is initially synthesized (Phelps and Myles, 1987). Although the nature of the specific targeting signals for the plasma membrane and secretory granule membrane is uncertain (Burgess and Kelly, 1987; Leube et al., 1989; Chung et al., 1989; Klausner, 1989), two general kinds of hypotheses can be proposed to explain PH-20 targeting. One hypothesis is that two distinct forms of PH-20 are encoded which differ in their primary sequence and thus in their targeting information, so that one is sent to the acrosomal membrane and the other to

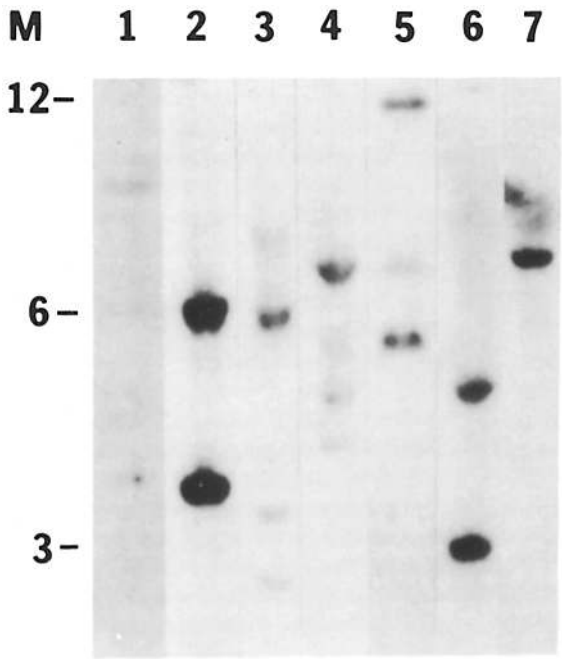


Figure 6. Southern blot of other species' genomic DNA. Size markers (*M*) are in kb. Lanes *1–7* contain genomic DNA (10 μg/lane). The source of DNA is (lane *1*) chicken; (lane *2*) guinea pig; (lane *3*) mouse; (lane *4*) rat; (lane *5*) hamster; (lane *6*) rabbit; and (lane *7*) human. The hamster DNA was digested with Pvu II and all others with Bgl II. The probe was labeled gpPH-20-2 and gpPH-20-3.

the plasma membrane. The alternative hypothesis supposes that only one form of PH-20 is encoded. Developmental changes during round spermatid differentiation could result in this single form being targeted first to the acrosomal membrane, where PH-20 is first detected (Phelps and Myles, 1987), and subsequently to the plasma membrane. Supporting evidence for this second type of hypothesis comes from the work of Braun et al. (1989). They found in transgenic mice that if the human growth hormone gene was expressed in early round spermatids, growth hormone was targeted to the acrosome; if the gene was expressed later, in condensing spermatids, the growth hormone was not targeted to the acrosome but to a different intracellular location. Thus, dependent upon the time of expression in spermatids, the product of a single gene was targeted first to the acrosome and second to another location.

All the data we have obtained are consistent with the hypothesis that only one form of PH-20 is encoded. Previous work indicated that there is two times as much PH-20 on the acrosomal membrane as on the plasma membrane of cauda epididymal sperm (Cowan et al., 1986). Thus, if two structurally different forms are made, both forms should be readily detectable. PH-20 protein, extracted from both membranes of cauda sperm and purified, migrates on SDS-PAGE as a single major band, M_r ~64. Lower molecular mass forms are found but experiments indicate that they are proteolytic fragments of 64 kD, as discussed above (Primakoff et al., 1988*a*; our unpublished results). The purified 64-kD form migrates as a single spot on two dimensional (2D) gel electrophoresis (Primakoff et al., 1988*a*). In our current results, Southern blots indicate there is one gene for PH-20 and Northern blots show one size message for PH-20. If there are two structurally distinct forms of PH-20, the 2D gel and Northern blot data show that the difference between them must be subtle. No sequence variants have been found in sequencing two clones coding for the 3′ half of the cDNA and three clones for the 5′ half. So the data obtained in this study have revealed only one coding form, and further investigation, particularly looking for sequence variants, will be required to search for a hypothetical second form.

Homology of PH-20 with Other Proteins

Given the primary sequence and information on the structural organization of PH-20 obtained in this study, we can ask if PH-20 is related to sperm proteins proposed to play a role in sperm–zona adhesion in other species. Sperm–zona adhesion has been extensively studied using mouse gametes. In the mouse, only acrosome-intact sperm will initiate binding to the zona (Saling et al., 1979; Saling and Storey, 1979; Florman and Storey, 1982; Bleil and Wassarman, 1983; Storey et al., 1984; Lee and Storey, 1985), although after these sperm undergo the acrosome reaction, the resulting acrosome-reacted sperm bind to the zona. In guinea pig, as in Chinese hamster, rabbit, and human, both acrosome-intact and acrosome-reacted sperm will initiate binding to the zona (Yanagimachi et al., 1983; Kuzan et al., 1984; Myles et al., 1987; Morales et al., 1989). These apparent phenomenological differences give no clue as to which adhesion proteins are involved. For instance, it is possible that a particular sperm adhesion protein might be located exclusively on the plasma membrane and act in acrosome-intact sperm adhesion in the mouse and be located primarily on the acrosomal membrane and act in acrosome-reacted sperm adhesion in the guinea pig. Thus, we have asked if the PH-20 protein of guinea pig is related to any known or proposed sperm adhesion proteins from other species.

Bindin (M_r 24) from sea urchin sperm is a sperm adhesion protein that binds to the coat of sea urchin eggs (Vacquier and Moy, 1977). Cloning of bindin cDNA showed that bindin is initially made as a large (51 kD) precursor which is processed to the mature 24-kD form. The bindin sequence shows no similarities to PH-20; in particular, mature bindin has no cysteines, no membrane anchor, and its hydropathy plot does not resemble PH-20 (Gao et al., 1986).

Mouse sperm have on their plasma membrane a beta-1,4-galactosyl transferase (M_r 60), which is an adhesion protein that functions in acrosome-intact sperm binding to the zona (Shur and Hall, 1982; Lopez et al., 1985; Shur and Neely, 1988; Shur, 1989*a*). cDNAs for murine beta-1,4-galactosyl transferase have been cloned and sequenced. Two slightly different cDNAs were found that code for two proteins differing in their NH_2-terminal domain (11 or 24 amino acids in the NH_2-terminal domain, Shaper et al., 1988). Although the cDNA clones were isolated from a library generated from mRNA from a mouse tissue culture line (Shaper et al., 1988), the cloned cDNAs hybridize to polyA$^+$ RNA from mouse testis and thus presumably code for the mouse sperm surface beta-1,4-galactosyl transferase (Shaper et al., 1990). The cDNAs for murine beta-1,4-galactosyl transferase reveal a protein that has no similarity to PH-20, either in its sequence or in its domain organization. The murine beta-1,4-galactosyl transferase has a short amino terminus (11 or 24

amino acids) that extends into the cytoplasm, a transmembrane domain and a large carboxy terminus (355 amino acids) containing the active site oriented extracellularly (Shaper et al., 1988). This basic domain structure (very short cytoplasmic NH_2-terminal domain, transmembrane domain, large extracellular COOH-terminal domain) has been found in an alpha-1,3-galactosyl transferase and in all other glycosyl transferases for which cDNAs have been isolated (Weinstein et al., 1987; Shaper et al., 1988; Joziasse et al., 1989; Russo et al., 1990). Furthermore, these glycosyl transferases have common features in their mRNAs, including an unusually long 3′ untranslated region (2.2–2.7 kb) which PH-20 does not share. Thus, we conclude that PH-20 is not related to the mouse sperm adhesion protein beta-1,4-galactosyl transferase and is probably not any other kind of glycosyl transferase.

A surface protein has been found on rat sperm which is a galactose receptor (RTG-r, M_r 54), is immunologically related to the hepatic asialoglycoprotein receptor, and has been hypothesized to have a role in rat sperm adhesion to the zona (Abdullah and Kierszenbaum, 1989). Although there is no structural information available about RTG-r on sperm, cDNA for the serologically related rat liver asialoglycoprotein receptor has been cloned and sequenced (Halberg et al., 1987). This cDNA shows the asialoglycoprotein receptor is a member of a family of animal lectin receptors which have a common type of carbohydrate-binding domain (Drickamer, 1988). Several recent reports show that adhesion proteins on various cell types belong to this family, contain a related carbohydrate-binding domain, and probably function by binding carbohydrate on their adhesion partners (Drickamer, 1988; Lasky et al., 1989; Siegelman et al., 1989; Tedder et al., 1989; Johnston et al., 1989; Bevilacqua et al., 1989). Examination of the PH-20 cDNA sequence shows that PH-20 does not contain this common carbohydrate binding domain and is not a member of this family of adhesion proteins. Further study of rat sperm RTG-r is necessary to find out if it does have a role in sperm adhesion to the zona and if it is a member of the family of animal lectin receptors. If RTG-r does prove to be a member of this family, then it is unlikely to be related to PH-20.

Several other sperm proteins with a proposed role in sperm–zona adhesion, i.e., the mouse 56-kD protein with affinity for ZP-3 (Bleil and Wassarman, 1990), mouse 95-kD (Leyton and Saling, 1989), rat 2B1 (Shalgi et al., 1990), boar AP_z (Peterson and Hunt, 1989), and rabbit RSA (O'Rand et al., 1988), do not yet have published sequence information. Comparison with the PH-20 sequence thus awaits more detailed information about these molecules.

In addition to asking if the PH-20 protein is related to other sperm proteins, we also asked if PH-20 is related to adhesion proteins found on other cell types or proteins with other activities. Computer searches revealed that PH-20 only has short sequence similarities with other known sequences. Some of these short PH-20 sequences show weak, possibly chance, similarities with bacterial or eukaryotic cytoplasmic proteins and have no apparent significance. Two sequence similarities of potential significance were found with the leukocyte integrin CD11c/CD18. CD11c/CD18 is a member of the integrin family of cell adhesion proteins found on myeloid cells, activated lymphocytes, neutrophils and monocytes (Corbi et al., 1987; Kishimoto et al., 1987; Hynes, 1987; Ruoslahti and Pierschbacher, 1987). In common with PH-20, it has a required role in cell–cell adhesion and it is found on both the plasma membrane and secretory granule membrane (Springer et al., 1984; Miller et al., 1987; Anderson and Springer, 1987). The CD11c/CD18 adhesion protein, like all integrins, is composed of two noncovalently associated subunits, termed alpha and beta, which are coded by two separate genes. The striking feature of the PH-20 sequence similarity is that PH-20 cDNA contains one short sequence similar to an alpha cDNA sequence and a second short sequence similar to a beta cDNA sequence. Specifically, PH-20 has a 31 residue sequence similar to a 31 residue sequence in alpha's extracellular domain and a 35 residue sequence similar to a 35 residue sequence in beta's extracellular domain. The regions of sequence similarity seem too short to indicate a common ancestral gene or evolutionary conservation of a protein domain involved in an adhesion function. Possibly, the two similar sequences in CD11c/CD18 alpha and beta and in PH-20 might be a case of convergent evolution involved in protein targeting or some other unknown function.

With the exception of this short, two region sequence similarity to CD11c/CD18, the significance of which is uncertain, computer searches revealed that PH-20 is a novel protein with no significant homology with other sequences in the data bases.

Presence of PH-20 in Other Species

We found, using cross-species Southern blots, that a homologue of the PH-20 gene is present in genomic DNA in a range of mammalian species and is also weakly detectable in chicken DNA. We have previously tested the cross-species reactivity of anti-PH-20 antibodies. Three mAbs to different epitopes of guinea pig PH-20 do not bind to sperm from rat, mouse, hamster, or human (Primakoff et al., 1983). Most anti-PH-20 polyclonal sera (raised against purified PH-20 protein in rabbits or guinea pigs, Primakoff et al., 1988*b*) are positive for binding to other species' sperm in some assays (immunofluorescence, radioimmune assay), but negative in other assays (Western blot, immunoprecipitation). The data suggest the possibility of a low affinity binding of anti-guinea pig PH-20 sera to other species' sperm, but alternative explanations are possible. Thus, the antibody data fail to provide a convincing answer about the presence of a PH-20 homologue in other species. In contrast, using guinea pig PH-20 cDNA probe and conditions of medium stringency, a strong signal was reproducibly obtained in Southern blots of other mammalian species' DNA. Work in progress shows that transcripts, hybridizing with guinea pig PH-20 cDNA probe, are expressed in the testis in other mammals (our unpublished observations). These results suggest that significant parts of the PH-20 gene are highly conserved but the parts coding for the most immunogenic regions of guinea pig PH-20 are not sufficiently conserved to yield strongly cross-reactive antisera.

Although mammalian sperm–zona adhesion is in most cases species specific, it has been found that genes for zona glycoproteins have been conserved during mammalian evolution (Ringuette et al., 1986; Ringuette et al., 1988; Lunsford et al., 1990; Moller et al., 1990). The conservation of the PH-20 gene suggests that species specificity of adhesion may arise because during evolution relevant sperm and egg

proteins change their binding domains (or other regions that can affect binding) while they maintain many of their other structural features.

In general, cloning of cDNAs for putative or demonstrated cell adhesion proteins has led to rapid progress in understanding their precise function and general biology (for reviews of recent progress on adhesion proteins of many cell types, see Damsky, 1989). Availability of cloned PH-20 cDNA will allow new approaches to (*a*) test directly PH-20's putative adhesion function in transfected cells or transgenic animals; (*b*) define the mechanism(s) that target PH-20 to two membranes; and (*c*) gain insight into the molecular basis of the species specificity of sperm–zona adhesion.

We would like to thank numerous people for advice on experimental protocols, particularly Mike Sussman, Hyun-Duck Nah, and Drs. Joel Pachter, Larry Klobutcher, Jurrien Dean, and Steve Padula. We thank Kathy Stone and Dr. Ken Williams for determining the sequences of tryptic peptides. We thank Dr. Nick Cross for doing immunofluoresence on human sperm with a rabbit anti-PH-20 polyclonal antiserum. We are particularly grateful to Dr. Russell Doolittle for doing homology searches on PH-20 and Dr. Kurt Drickamer for analyzing the sequence for its relationship to the animal lectin family.

This work was supported by National Institutes of Health grants HD-16580 (D. G. Myles) and HD-21989 (P. Primakoff), an American Cancer Society Faculty Research Award (D. G. Myles), and a Labor Foundation Fellowship (E. P. Carmichael).

Received for publication 20 July 1990 and in revised form 13 August 1990.

References

Abdullah, M., and A. L. Kierszenbaum. 1989. Identification of rat testis galactosyl receptor using antibodies to liver asialoglycoprotein receptor: purification and localization on surfaces of spermatogenic cells and sperm. *J. Cell Biol.* 108:367–375.

Anderson, D. C., and T. A. Springer. 1987. Leukocyte adhesion deficiency: an inherited defect in the Mac-1, LFA-1, and p150,95 glycoproteins. *Annu. Rev. Med.* 38:175–194.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1987. *In* Current Protocols in Molecular Biology. Vol. 1. Greene Publishing Associates, New York, NY.

Bevilacqua, M. P., S. Stengelin, M. A. Gimbrone, Jr., and B. Seed. 1989. Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils related to complement regulatory proteins and lectins. *Science (Wash. DC).* 243:1160–1165.

Bleil, J. D., and P. M. Wassarman. 1980. Mammalian sperm-egg interaction: identification of a glycoprotein in mouse egg zonae pellucidae possessing receptor activity for sperm. *Cell.* 20:873–882.

Bleil, J. D., and P. M. Wassarman. 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, J. D., and P. M. Wassarman. 1990. Identification of a ZP-3 binding protein an acrosome-intact mouse sperm by photoaffinity crosslinking. *Proc. Natl. Acad. Sci. USA.* 87:5563–5567.

Braun, R. E., J. J. Peschon, R. R. Behringer, R. L. Brinster, and R. D. Palmiter. 1989. Protamine 3′-untranslated sequences regulate temporal translational control and subcellular localization of growth hormone in spermatids of transgenic mice. *Genes & Dev.* 3:793–802.

Brümmendorf, T., J. M. Wolff, R. Frank, and F. G. Rathjen. 1989. Neural cell recognition molecule F11: homology with fibronectin type III and immunoglobulin type C domains. *Neuron.* 2:1351–1361.

Burgess, T. L., and R. B. Kelly. 1987. Constitutive and regulated secretion of proteins. *Annu. Rev. Cell Biol.* 3:243–294.

Camerini, D., S. P. James, I. Stamenkovic, and B. Seed. 1989. Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing receptor. *Nature (Lond.).* 342:78–82.

Chung, K.-N., P. Walter, G. W. Aponte, and H.-P. H. Moore. 1989. Molecular sorting in the secretory pathway. *Science (Wash. DC).* 243:192–197.

Corbi, A. L., L. J. Miller, K. O'Connor, R. S. Larson, and T. A. Springer. 1987. cDNA cloning and complete primary structure of the α subunit of a leukocyte adhesion glycoprotein, p150,95. *EMBO (Eur. Mol. Biol. Organ.) J.* 6: 4023–4028.

Cowan, A. E., P. Primakoff, and D. G. Myles. 1986. Sperm exocytosis increases the amount of PH-20 antigen on the surface of guinea pig sperm. *J. Cell Biol.* 103:1289–1297.

Damsky, C. H., editor. 1989. Cell-to-cell contact. *Curr. Opinion Cell Biol.* 1:881–952.

Drickamer, K. 1988. Two distinct classes of carbohydrate-recognition domains in animal lectins. *J. Biol. Chem.* 263:9557–9560.

Dustin, M. L., P. Selvaraj, R. J. Mattaliano, and T. A. Springer. 1987. Anchoring mechanisms of LFA-3 cell adhesion glycoprotein at membrane surface. *Nature (Lond.).* 329:846–848.

Ferguson, M. A., and A. F. Williams. 1988. Cell-surface anchoring of proteins via glycosyl-phosphatidylinositol structures. *Annu. Rev. Biochem.* 57:285–320.

Florman, H. M., and B. T. Storey. 1982. Mouse gamete interactions: the zona pellucida is the site of the acrosome reaction leading to fertilization *in vitro*. *Dev. Biol.* 91:121–130.

Frohman, M. A., M. K. Dush, and G. R. Martin. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA.* 85:8998–9002.

Gao, B., L. E. Klein, R. J. Britten, and E. H. Davidson. 1986. Sequence of mRNA coding for bindin, a species-specific sea urchin sperm protein required for fertilization. *Proc. Natl. Acad. Sci. USA.* 83:8634–8638.

Geng, J.-G., M. P. Bevilacqua, K. L. Moore, T. M. McIntyre, S. M. Prescott, J. M. Kim, G. A. Bliss, G. A. Zimmerman, and R. P. McEver. 1990. Rapid neutrophil adhesion to activated endothelium mediated by GMP-140. *Nature (Lond.).* 343:757–760.

Gennarini, G., G. Cibelli, G. Rougon, M. Mattei, and C. Goridis. 1989. The mouse neuronal cell surface protein F3: a phosphatidylinositol-anchored member of the immunoglobulin superfamily related to chicken contactin. *J. Cell Biol.* 109:775–788.

Gubler, U., and B. J. Hoffmann. 1983. A simple and very efficient method for generating cDNA libraries. *Gene.* 25:263–269.

Halberg, D. F., R. E. Wager, D. C. Farrell, J. Hildreth IV, M. S. Quesenberry, J. A. Loeb, E. C. Holland, and K. Drickamer. 1987. Major and minor forms of the rat liver asialoglycoprotein receptor are independent galactose-binding proteins. *J. Biol. Chem.* 262:9828–9838.

He, H.-T., J. Barbet, J.-C. Chaix, and C. Goridis. 1986. Phosphatidylinositol is involved in the membrane attachment of NCAM-120, the smallest component of the neural cell adhesion molecule. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2489–2494.

Hemperly, J. J., G. M. Edelman, and B. A. Cunningham. 1986. cDNA clones of the neural cell adhesion molecule (N-CAM) lacking a membrane-spanning region consistent with evidence for membrane attachment via a phosphatidylinositol intermediate. *Proc. Natl. Acad. Sci. USA.* 83:9822–9826.

Hottori, R., K. K. Hamilton, R. D. Fugates, R. P. McEver, and P. J. Sims. 1989. Stimulated secretion of endothelial von Willebrand factor is accompanied by rapid redistribution to the cell surface of the intracellular granule membrane protein GMP-140. *J. Biol. Chem.* 264:7768–7771.

Hsu-Lin, S.-C., C. L. Berman, B. C. Furie, D. August, and B. Furie. 1984. A platelet membrane protein expressed during platelet activation and secretion. *J. Biol. Chem.* 259:9121–9126.

Hynes, R. O. 1987. Integrins: a family of cell surface receptors. *Cell.* 48:549–554.

Johnston, G. I., R. G. Cook, and R. P. McEver. 1989. Cloning of GMP-140, a granule membrane protein of platelets and endothelium: sequence similarity to proteins involved in cell adhesion and inflammation. *Cell.* 56:1033–1044.

Jones, R., R. Shalgi, J. Hoyland, and D. M. Phillips. 1990. Topographical rearrangement of a plasma membrane antigen during capacitation of rat spermatozoa *in vitro*. *Dev. Biol.* 139:349–362.

Joziasse, D. H., J. H. Shaper, D. H. van den Eijnden, A. J. Van Tunen, and N. L. Shaper. 1989. Bovine α1,3-galactosyltransferase: isolation and characterization of a cDNA clone. *J. Biol. Chem.* 264:14290–14297.

Kishimoto, T. K., K. O'Connor, A. Lee, T. M. Roberts, and T. A. Springer. 1987. Cloning of the β subunit of the leukocyte adhesion proteins: homology to an extracellular matrix receptor defines a novel supergene family. *Cell.* 48:681–690.

Kishimoto, T. K., M. A. Jutlia, E. L. Berg, and E. C. Butcher. 1989. Neutrophil Mac-1 and MEL-14 adhesion proteins inversely regulated by chemotactic factors. *Science (Wash. DC).* 245:1238–1241.

Klausner, R. D. 1989. Sorting and traffic in the central vacuolar system. *Cell.* 57:703–706.

Kuzan, F. B., A. D. Fleming, and G. E. Seidel, Jr. 1984. Successful fertilization *in vitro* of fresh intact oocytes by perivitelline (acrosome-reacted) spermatozoa of the rabbit. *Fertil. Steril.* 41:766–770.

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

Lasky, L. A., M. S. Singer, T. A. Yednock, D. Dowbenko, C. Fennie, H. Rodriguez, T. Nguyen, S. Stachel, and S. D. Rosen. 1989. Cloning of a lymphocyte homing receptor reveals a lectin domain. *Cell.* 56:1045–1055.

Lee, M. A., and B. T. Storey. 1985. Evidence for plasma membrane impermeability to small ions in acrosome-intact mouse spermatozoa bound to mouse zonae pellucidae, using an aminoacridine fluorescent pH probe: time course of the zona-induced acrosome reaction monitored by both chlortetracycline and pH probe fluorescence. *Biol. Reprod.* 33:235–246.

Leube, R. E., B. Wiedenmann, and W. W. Franke. 1989. Topogenesis and sorting of synaptophysin: synthesis of a synaptic vesicle protein from a gene transfected into nonneuroendocrine cells. *Cell.* 59:433–446.

Leyton, L., and P. Saling. 1989. 95 kd sperm proteins bind ZP3 and serve as

tyrosine kinase substrates in response to zona binding. *Cell.* 57:1123–1130.
Lopez, L. C., E. M. Bayna, D. Litoff, N. L. Shaper, J. H. Shaper, and B. D. Shur. 1985. Receptor function of mouse sperm surface galactosyltransferase during fertilization. *J. Cell Biol.* 101:1501–1510.
Low, M. G. 1989. Glycosyl-phosphatidylinositol: a versatile anchor for cell surface proteins. *FASEB (Fed. Am. Soc. Exp. Biol.) J.* 3:1600–1608.
Low, M. G., and A. R. Saltiel. 1988. Structural and functional roles of glycosylphosphatidylinositol in membranes. *Science (Wash. DC).* 239: 268–275.
Lunsford, R. D., N. A. Jenkins, C. A. Kozak, L. Liang, C. M. Silan, N. G. Copeland, and J. Dean. 1990. Genomic mapping of murine Zp-2 and Zp-3, two oocyte-specific loci encoding zona pellucida proteins. *Genomics.* 6:184–187.
Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. 545 pp.
McEver, R. P., and M. N. Martin. 1984. A monoclonal antibody to a membrane glycoprotein binds only to activated platelets. *J. Biol. Chem.* 259:9799–9804.
Mierendorf, R. C., C. Percy, and R. A. Young. 1987. Gene isolation by screening 2gt11 libraries with antibodies. *In* Methods Enzymol. 152:458–464.
Miller, L. J., D. F. Bainton, N. Borregaard, and T. A. Springer. 1987. Stimulated mobilization of monocyte-Mac-1 and p150,95 adhesion proteins from an intracellular vesicular compartment to the cell surface. *J. Clin. Invest.* 80:535–544.
Moller, C. C., J. D. Bleil, R. A. Kinloch, and P. M. Wassarman. 1990. Structural and functional relationships between mouse and hamster zona pellucida glycoproteins. *Dev. Biol.* 137:276–286.
Morales, P., N. L. Cross, J. W. Overstreet, and F. W. Hanson. 1989. Acrosome intact and acrosome-reacted human sperm can initiate binding to the zona pellucida. *Dev. Biol.* 133:385–392.
Myles, D. G., H. Hyatt, and P. Primakoff. 1987. Binding of both acrosome-intact and acrosome-reacted guinea pig sperm to the zona pellucida during *in vitro* fertilization. *Dev. Biol.* 121:559–567.
O'Rand, M. G., E. E. Widgren, and S. J. Fisher. 1988. Characterization of the rabbit sperm membrane autoantigen, RSA, as a lectin-like zona binding protein. *Dev. Biol.* 129:231–240.
Peterson, R. N., and W. P. Hunt. 1989. Identification, isolation, and properties of a plasma membrane protein involved in the adhesion of boar sperm to the porcine zona pellucida. *Gamete Res.* 23:103–118.
Phelps, B. M., and D. G. Myles. 1987. The guinea pig sperm plasma membrane protein, PH-20, reaches the surface via two transport pathways and becomes localized to a domain after an initial uniform distribution. *Dev. Biol.* 123:63–72.
Phelps, B. M., P. Primakoff, D. E. Koppel, M. G. Low, and D. G. Myles. 1988. Restricted lateral diffusion of PH-20, a PI-anchored sperm membrane protein. *Science (Wash. DC).* 240:1780–1782.
Phelps, B. M., D. E. Koppel, P. Primakoff, and D. G. Myles. 1990. Evidence that proteolysis of the surface is an initial step in the mechanism of formation of sperm cell surface domains. *J. Cell Biol.* 111:1839–1848.
Primakoff, P., M. R. Heaton, and D. G. Myles. 1983. Species specificity of hybridoma antibodies to surface antigens of guinea pig sperm. *Gamete Res.* 8:385–394.
Primakoff, P., H. Hyatt, and D. G. Myles. 1985. A role for the migrating sperm surface antigen PH-20 in guinea pig sperm binding to the egg zona pellucida. *J. Cell Biol.* 101:2239–2244.
Primakoff, P., A. Cowan, H. Hyatt, J. Tredick-Kline, and D. G. Myles. 1988*a*. Purification of the guinea pig sperm PH-20 antigen and detection of a site-specific endoproteolytic activity in sperm preparations that cleaves PH-20 into two disulfide-linked fragments. *Biol. Reprod.* 38:921–934.
Primakoff, P., W. Lathrop, L. Woolman, A. Cowan, and D. Myles. 1988*b*. Fully effective contraception in male and female guinea pigs immunized with the sperm protein PH-20. *Nature (Lond.).* 335:543–546.
Ringuette, M. J., D. A. Sobieski, S. M. Chamow, and J. Dean. 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.
Ringuette, M. J., M. E. Chamberlin, A. W. Baur, D. A. Sobieski, and J. Dean. 1988. Molecular analysis of cDNA coding for ZP3, a sperm binding protein of the mouse zona pellucida. *Dev. Biol.* 127:287–295.
Ruoslahti, E., and M. D. Pierschbacher. 1987. New perspectives in cell adhesion: RGD and integrins. *Science (Wash. DC).* 238:491–497.
Russo, R. N., N. L. Shaper, and J. H. Shaper. 1990. Bovine β1, 4-galactosyltransferase: two sets of mRNA transcripts encode two forms of the protein with different amino-terminal domains. *J. Biol. Chem.* 265:3324–3331.
Saling, P. M. 1989. Mammalian sperm interaction with extracellular matrices of the egg. *In* Oxford Reviews of Reproductive Biology. Vol. 11. S. R. Milligan, editor. Oxford University Press, Oxford, England. 340–388.
Saling, P. M., and B. T. Storey. 1979. Chlortetracycline as a fluorescent probe for the mouse sperm acrosome reaction. *J. Cell Biol.* 83:544–555.
Saling, P. M., J. Sowinski, and B. T. Storey. 1979. An ultrastructural study of epididymal mouse spermatozoa binding to zonae pellucidae *in vitro*: sequential relationship to the acrosome reaction. *J. Exp. Zool.* 209:229–238.
Schmell, E. D., and B. J. Gulyas. 1980. Mammalian sperm-egg-recognition and binding in vitro. I. Specificity of sperm interactions with live and fixed eggs in homologous and heterologous inseminations of hamster, mouse, and guinea pig oocytes. *Biol Reprod.* 23:1075–1085.
Seed, B. 1987. An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. *Nature (Lond.).* 329:840–845.
Shalgi, R., A. Matityahu, S. J. Gaunt, and R. Jones. 1990. Antigens on rat spermatozoa with a potential role in fertilization. *Mol. Reprod. Dev.* 25:286–296.
Shaper, N. L., G. F. Hollis, J. G. Douglas, I. R. Kirsch, and J. H. Shaper. 1988. Characterization of the full length cDNA for murine β-1,4-galactosyltransferase. *J. Biol. Chem.* 263:10420–10427.
Shaper, N. L., W. W. Wright, and J. H. Shaper. 1990. Murine beta 1,4-galactosyl-transferase: both the amounts and structure of the mRNA are regulated during spermatogenesis. *Proc. Natl. Acad. Sci. USA.* 87(2): 791–795.
Shur, B. D. 1989*a*. Galactosyltransferase as a recognition molecule during fertilization and development. *In* The Molecular Biology of Fertilization. H. Schatten and G. Schatten, editors. Academic Press Inc., Orlando, FL. 37–71.
Shur, B. D. 1989*b*. Glycoconjugates as mediators of cellular interactions during development. *Current Opinion Cell Biol.* 1:905–912.
Shur, B. D., and N. G. Hall. 1982. A role for mouse sperm surface galactosyltransferase in sperm binding to the egg zona pellucida. *J. Cell Biol.* 95:574–579.
Shur, B. D., and C. A. Neely. 1988. Plasma membrane association, purification, and partial characterization of mouse sperm β1,4-galactosyltransferase. *J. Biol. Chem.* 263:17706–17714.
Siegelman, M. H., M. van de Rijn, and W. L. Weissman. 1989. Mouse lymph node homing receptor cDNA clone encodes a glycoprotein revealing tandem interaction domains. *Science (Wash. DC).* 243:1165–1172.
Springer, T. A., W. S. Thompson, L. J. Miller, F. C. Schmalstieg, and D. C. Anderson. 1984. Inherited deficiency of the Mac-1, LFA-1, p150,95 glycoprotein family and its molecular basis. *J. Exp. Med.* 160:1901–1918.
Stone, K. L., M. B. Lopresti, J. M. Crawford, R. DeAngelis, and K. R. Williams. 1989. Enzymatic digestion of proteins and HPLC peptide isolation. *In* A Practical Guide to Protein and Peptide Purification for Microsequencing. P. Matsudaira, editor. Academic Press Inc., Orlando, FL. 31–47.
Storey, B. Y., M. A. Lee, C. Muller, C. R. Ward, and D. G. Wirtshafter. 1984. Binding of mouse spermatozoa to the zonae pellucidae of mouse eggs in cumulus: evidence that the acrosomes remain substantially intact. *Biol. Reprod.* 31:1119–1128.
Tedder, T. E., C. M. Isaacs, T. J. Ernst, G. D. Demetri, D. A. Adler, and C. M. Disteche. 1989. Isolation and chromosomal localization of cDNAs encoding a novel human lymphocyte cell surface molecule, LAM-1. *J. Exp. Med.* 170:123–133.
Todd, R. F., III, M. A. Arnaout, R. E. Rosin, C. A. Crowley, W. A. Peters, and B. M. Babior. 1984. Subcellular localization of the large subunit of Mo1 (Mo1α; formerly gp 110), a surface glycoprotein associated with neutrophil adhesion. *J. Clin. Invest.* 74:1280–1290.
Vacquier, V. T., and G. W. Moy. 1977. Isolation of bindin: the protein responsible for adhesion of sperm to sea urchin eggs. *Proc. Natl. Acad. Sci. USA.* 74:2456–2460.
von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14:4683–4690.
von Heijne, G. 1987. Sequence Analysis in Molecular Biology. 113–118.
von Heijne, G. 1990. The signal peptide. *J. Membr. Biol.* 115:195–201.
Weinstein, J., E. U. Lee, K. McEntee, P.-H. Lai, and J. Paulson. 1987. Primary structure of β-galactoside α2,6-sialyltransferase. *J. Biol. Chem.* 262:17735–17743.
Yanagimachi, R. 1981. Mechanisms of fertilization in mammals. *In* Fertilization and Embryonic Development *In Vitro*. L. Mastroianni, Jr. and J. D. Biggers, editors. Plenum Publishing Corp., New York. 81–182.
Yanagimachi, R., Y. Kamiguchi, S. Sugawara, and K. Mikamo. 1983. Gametes and fertilization in the Chinese hamster. *Gamete Res.* 8:97–117.
Yanigimachi, R. 1988. Mammalian fertilization. *In* Physiology of Reproduction. Vol. 1. E. Knobil and J. D. Neill, editors. Raven Press, New York. 135–185.