A GPI-anchored Sea Urchin Sperm Membrane Protein Containing EGF Domains is Related to Human Uromodulin

Lisa M. Mendoza*, David Nishioka† and Victor D. Vacquier*

*Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093-0202; and †Department of Biology, Georgetown University, Washington, DC 20057

Abstract. An Mr 63-kD sea urchin sperm flagellar membrane protein has been previously implicated as a possible receptor for egg jelly ligand(s) that trigger the sperm acrosome reaction (AR). The cDNA and deduced amino acid sequences of the 63-kD protein are presented. The open reading frame codes for a protein of 470 amino acids which contains a putative signal sequence of 25 residues. Western blots using antibodies to two synthetic peptides confirm the sequence to be that of the 63-kD protein. The mRNA is ~2,300 bases in length and the gene appears to be single copy. The protein is released from sperm membrane vesicles by treatment with phosphatidylinositol-specific phospholipase C, showing that it is anchored to the flagellar membrane by glycosylphosphatidyl inositol (GPI). Although we cannot demonstrate involvement of the 63-kD protein in the AR, it is of potential interest because it shares significant similarity with the developmentally expressed proteins crumbs, notch and xotch as well as human uromodulin over a region that includes two separate EGF repeats.

Sperm-egg interactions during fertilization provide important models for studying such basic cellular phenomena as chemotaxis (Ward et al., 1985; Rait et al., 1991), cell-cell recognition and adhesion (Wasserman, 1990; Foltz and Lennarz, 1992), membrane fusion (Hong and Vacquier, 1986; Blobel et al., 1992; White, 1992), oxidative stress (Shapiro, 1990), and ionic (Epel, 1990) and second messenger (Garbers, 1989)-mediated cellular activation. When exposed to the extracellular jelly layer of the unfertilized egg, sea urchin sperm undergo the acrosome reaction (AR) in which the acrosomal vesicle is exocytosed and an acrosomal process of filamentous actin is extended from the tip of the sperm head (Dan, 1967). The AR is an absolute prerequisite for fertilization; it is induced by the opening of ligand-gated ion channels, resulting in the net influx of Ca²⁺ and Na⁺, and the net efflux of H⁺ and K⁺ (Vacquier, 1986; Babcock et al., 1992; Gonzalez-Martinez et al., 1992).

mAb J18/29 induces the AR in sea urchin sperm (Trimmer et al., 1987). Based on reaction with this mAb, sperm membrane proteins of approximate Mr, 320 kD, 210 kD, 170 kD and 63 kD have been implicated as potential receptors for the egg jelly peptide speract that activates sperm respiration and motility (Harumi et al., 1991). Another mAb, J17/30, reacts exclusively with the 63-kD protein and localizes it to the sperm flagellum and midpiece (Nishioka et al., 1987).

To gain a deeper understanding of the sperm membrane proteins mediating the AR, we utilized molecular techniques to characterize the 63-kD protein. Here we report the cDNA and deduced amino acid sequences of the 63-kD protein of Strongylocentrotus purpuratus sperm. We show that the protein is anchored to the membrane by glycosylphosphatidylinositol (GPI) and that it shares significant similarity with proteins of the EGF superfamily, including human uromodulin (Hession et al., 1987), and the developmentally regulated proteins crumbs (Tepass et al., 1990), notch (Wharton et al., 1985) and xotch (Coffman et al., 1990).

Materials and Methods

Cloning and Sequencing the cDNA Encoding the 63-kD Protein

A 1.9-kb cDNA coding for the 63-kD protein was isolated from a λgt 11 sea urchin (S. purpuratus) testis library using mAb J17/30 (Nishioka et al., 1987; Trimmer et al., 1987) after standard procedures for antibody screening (Young and Davis, 1983). The Eco RI insert was cloned into pBluescript (Stratagene Corp., La Jolla, CA) and single stranded DNA (M13) was isolated. The sequences of both strands were determined by the dideoxy chain termination method using Sequenase 2.0 (United States Biochemical,
Cleveland, OH). This clone was then used to isolate several overlapping cDNAs from an *S. purpuratus* testes cDNA XZAP II library (Stratagene Corp.). Hybridizations were at 65°C in 6× SSPE (1× SSPE: 150 mM NaCl, 1 mM EDTA, 10 mM NaH_2PO_4 pH 7.4), 5× Denhardt's (50× Denhardt's: 1% ficol, 1% polyvinylpyrrolidone, 1% BSA), 0.5% SDS and 100 μg/ml yeast tRNA (Sigma Chemical Co., St. Louis, MO) overnight. Initial washes were at 23°C for 15 min (2×) in 1× SSC, 0.1% SDS with a final wash at 68°C in 1× SSC, 0.1% SDS for 1 h.

**Analysis of Sequences**

The deduced amino acid sequence of the 63-kD cDNA was compared to all protein libraries available on the Pearson FAST88 package using Fasta (Pearson, 1990). An optimal alignment of homologous sequences was derived using the programs RELATE and ALIGN from the Protein Identification Resource (National Biomedical Research Foundation, Washington, DC). A hydrophathy profile of the deduced amino acid sequence was obtained using TUREASE on the Pearson FAST88 package using the Kyte-Doolittle algorithm (Kyte and Doolittle, 1982). All computer programs used were accessed on the VAX/VMS system (Smith, 1988).

**Peptide Antigen/Antibody Production and Immunoblotting**

Peptides 1 and 2 were synthesized based on the deduced amino acid sequence of the 63-kD protein from residues 237-248 and 382-393 (Fig. 1; underlines). Both peptides had an additional lysine at the amino terminus and were coupled to BSA by the glutaraldehyde method (Doolittle, 1987). The peptide antigens (BSA-peptide) were used to generate antisera in rabbits (Cocalico Biologicals, Reamstown, PA). Peptide-specific IgG was purified by affinity chromatography using an Immunopure Ag/Ab Immunobilization Kit No. 1 (Pierce Chemical Co., Rockford, IL). First, BSA-specific IgG was removed on a BSA affinity column, followed by purification of peptide-specific IgG on a BSA-peptide column. Immunoblotting of sperm membrane proteins isolated by the method of Podell et al. (1984) with peptide-specific IgG, and with mAb J17/30 IgG was detected with an alkaline phosphatase-conjugated secondary antibody (Calbiochem Corp., La Jolla, CA).

**Southern Blot Analysis**

Sperm DNA (5 μg) was digested with 20 U of *Pst I*, 15 U of *Bst XI*, 20 U of *Xba I*, and 10 U of *Nde I*, and electrophoresed in a 0.6% agarose gel. The DNA was transferred to a nylon membrane (Schleicher and Schuell, Inc., Keene, NH) and the blot probed with pl29-E, a 1,550-bp 63-103 eDNA fragment that was random-prime labeled with [α-^32P]dCTP (American Type Culture Collection, Rockville, MD). The DNA was transferred to a nylon membrane (Schieicher and Schuell, KGaA, Dassel, Germany) and the blot probed with pl29-E, a 1,550-bp 63-103 eDNA fragment that was random-prime labeled with [α-^32P]dCTP (American Type Culture Collection, Rockville, MD). The DNA was transferred to a nylon membrane (Schieicher and Schuell, KGaA, Dassel, Germany) and the blot probed with pl29-E, a 1,550-bp 63-103 eDNA fragment that was random-prime labeled with [α-^32P]dCTP (American Type Culture Collection, Rockville, MD). The DNA was transferred to a nylon membrane (Schieicher and Schuell, KGaA, Dassel, Germany) and the blot probed with pl29-E, a 1,550-bp 63-103 eDNA fragment that was random-prime labeled with [α-^32P]dCTP (American Type Culture Collection, Rockville, MD). The DNA was transferred to a nylon membrane (Schieicher and Schuell, KGaA, Dassel, Germany) and the blot probed with pl29-E, a 1,550-bp 63-103 eDNA fragment that was random-prime labeled with [α-^32P]dCTP (American Type Culture Collection, Rockville, MD).

**Peptide Specific IgG**

The peptide antigens (BSA-peptide) were used to generate antisera in rabbits (Cocalico Biologicals, Reamstown, PA). Peptide-specific IgG was purified by affinity chromatography using an Immunopure Ag/Ab Immunobilization Kit No. 1 (Pierce Chemical Co., Rockford, IL). First, BSA-specific IgG was removed on a BSA affinity column, followed by purification of peptide-specific IgG on a BSA-peptide column. Immunoblotting of sperm membrane proteins isolated by the method of Podell et al. (1984) with peptide-specific IgG, and with mAb J17/30 IgG was detected with an alkaline phosphatase-conjugated secondary antibody (Calbiochem Corp., La Jolla, CA).

**Figure 1.** The cDNA and deduced amino acid sequences of the 63-kD protein. Numbers on the right margin indicate either amino acid position where the putative start methionine is designated 1, or nucleotide sequence position. The numbers in each row correspond to the sequence position of the last amino acid/nucleotide in that row. A signal sequence of 25 amino acids (Von Heijne, 1986) is followed by the putative amino-terminal amino acid, glutamine (Δ). The mature sequence contains five sites for potential N-linked glycosylation (Δ). Synthetic peptide 1 is underlined with a solid line and peptide 2 with a dashed line. The stop codon (*) yields an open reading frame of 470 amino acids followed by a 3' untranslated region of 641 nucleotides containing three modified signal sequences for polyadenylation (AAATA). Glycine 446 (Δ) is tentatively assigned as the site of GPI attachment. These sequence data are available from EMBL/GenBank/DDBJ under accession number M99584.
Figure 4. Protein immunoblots with J17/30 and synthetic peptide-specific antibodies. Lanes A and B are sperm membrane vesicles (SMVs) from S. purpuratus (1.0 μg and 5.0 μg) reacted with the 63-kD-specific mAb J17/30 (2.0 μg/ml). Lanes C-E contain 20 μg of SMVs. Lane C was reacted with peptide 1 IgG at 2.5 μg/ml and lanes D and E were reacted with 2.5 μg/ml of peptide 2 IgG. Lane E was overdeveloped. The asterisk (*) denotes Mr 63 kD.

Figure 2. Northern blot analysis of S. purpuratus testis. Various amounts of poly A+ RNA were electrophoresed on a formaldehyde-denatured gel and blotted onto a nylon membrane, hybridized with clone pL29-E at 7 × 10^6 cpm/μg and washed at high stringency followed by autoradiography. RNA load (lane A) 22 μg; (lane B) 11 μg; (lane C) 5.5 μg; (lane D) 2.8 μg. Kilobase size markers are on the right.

To confirm that the size of the mRNA coding for the 63-kD protein corresponds to that of the cDNA presented in Fig. 1, Northern blot analysis of testicular poly A+ RNA was performed. Fig. 2 shows that 32P-labeled pL29-E, a 1,550 nucleotide fragment of the 63-kD protein cDNA, hybridizes to a single mRNA of ~2,300 nucleotides, quite close to 2,180 nucleotides presented in Fig. 1.

Southern blot analysis was performed with S. purpuratus sperm genomic DNA to determine the copy number of the 63-kD gene. Restriction enzyme digested genomic DNA from sperm cells was hybridized with 32P-labeled pL29-E (Fig. 3). Neither Nde I nor Xba I restriction sites are present within the pL29-E sequence. Fig. 3 shows that 32P-pL29-E hybridizes with a single band in both Nde I (N; at 21 kb) and Xba I (X; at 7 kb) digested genomic DNA, supporting the conclusion that the 63-kD protein is encoded by a single copy gene. Furthermore, both Bst XI and Pst I restriction sites are present at one site within the 1,550-bp pL29-E cDNA sequence. As expected for a single copy gene, two bands (at 17 and 4 kb) hybridize with 32P-pL29-E in the Pst I (P) digested genomic DNA. Hybridization of 32P-pL29-E to the Bst XI (B) digested genomic DNA, however, is more difficult to interpret and suggests the presence of introns within the gene.

To confirm that the deduced amino acid sequence was that of the 63-kD sperm membrane protein recognized by mAb J17/30, rabbit antibodies were prepared against two synthetic peptides made to distinct regions of the deduced sequence (underlines in Fig. 1; shaded zones in Fig. 5). On Western immunoblots of sperm membrane vesicles (SMVs), both peptide antisera reacted specifically with an antigen of Mr 63 kD, which comigrated with the sperm membrane protein recognized by mAb J17/30 (Fig. 4). We conclude that the deduced amino acid sequence is that of the 63-kD protein. Neither the two peptide antibodies nor mAb J17/30 induced sperm cells to undergo the AR, and none of the antibodies were capable of inhibiting fertilization or the egg jelly-induced AR.

Figure 3. Southern blot analysis. Each lane contains 5 μg of genomic DNA from sperm cells digested with various restriction enzymes (Nde I, N; Xba I, X; Bst XI, B; Pst I, P) as well as undigested (U) DNA. The blot was hybridized with clone pL29-E. Size markers are shown on the right in kilobases.

Figure 5. Hydropathy profile of the 63-kD deduced amino acid sequence. Hydrophobicity plotted against amino acid position was obtained with the program TGREASE. Hydrophobic index is on the y-axis in arbitrary units. A hydrophobic stretch at the extreme NH2-terminal end is designated as the signal sequence (SS). A single membrane spanning (MS) domain at the extreme COOH-terminal end is indicated. The regions of the deduced sequence from which peptides 1 and 2 were designed are shaded.
observed Mr of 63 kD. The 63 kD protein is not released.

Two regions of the 63-kD protein’s deduced amino acid sequence containing EGF-like domains are homologous to several proteins in GenBank. The first region spans ~38 residues, from amino acid position 43 to 81. The second region spans 113 residues, from amino acid position 202 to 315. The proteins identified from the GenBank search include the sea urchin EGF homolog (Hursh et al., 1987), uromodulin (also known as the Tamm-Horsfall protein; Hession et al., 1987) from human, crumbs (Tepass et al., 1990) and notch (Wharton et al., 1985) from Drosophila, and xotch (Coffman et al., 1990), the Xenopus homolog of notch.

To determine the significance of the similarity between these distantly related proteins, and to obtain optimal alignments between them and the 63-kD sea urchin sperm protein, the program ALIGN (Dayhoff et al., 1983) was used. EGF 1 contains six cysteine residues that are conserved between the 63-kD protein and the other EGF-like proteins (Fig. 7 a). The second region of similarity (Fig. 7 b) includes the two remaining EGF-like domains in the 63-kD protein sequence (EGF 2 and EGF 3). This region is characterized by the alignment of 11-13 cysteine residues with the 63-kD protein over 102-113 amino acids. Table 1 lists pairwise comparisons of the percent amino acid identity for these domains (above the diagonals) and the alignment score in SD units greater than random (below the diagonals) determined by the program ALIGN. Table 1, a and b, correspond to the regions presented in Fig. 7, a and b. Both percent identity and alignment scores presented in Table 1 include only the regions presented in Fig. 7, not the entire protein sequence.

The most significant homology is between the 63-kD protein and uromodulin, a glycoprotein expressed in human kidney on the thick ascending limb of the loop of Henle and Muchmore, 1990). The alignment score is 10 SD U above random (Table 1 b) which indicates that the probability of this alignment occurring by chance is <1 in 10^3. The least significant alignment score is 4.0 SD units between the 63-kD protein and xotch over the region that includes EGF 2 and 3 (Table 1 b). The probability of this occurring by chance is <1 in 10^4. These data support the conclusion that the sea urchin sperm 63-kD protein is a member of the EGF superfamily.

Homology to the EGF Superfamily

Two regions of the 63-kD protein’s deduced amino acid sequence containing EGF-like domains are homologous to several proteins in GenBank. The first region spans ~38 residues, from amino acid position 43 to 81. The second region spans 113 residues, from amino acid position 202 to 315. The proteins identified from the GenBank search in-
a. EGF 1

su63  DPGASW  DFTAST  78
suEGF  DPGASW  DFTAST  77
dCRU  DPGASW  DFTAST  77
xXOT  DPGASW  DFTAST  77

suEGF  DFTAST  77

b. EGF 2

Figure 7. Optimal alignments of sequences homologous to the 63-kD protein were derived using the programs RELATE and ALIGN. The matrix used for the alignment was PAM250; matrix bias was 6; and gap penalty was 12. Dots (-) indicate identical residues and dashes (-) are inserted for alignment. Conserved cysteine residues are shaded. The protein sequences include the sea urchin 63-kD protein (su63), the sea urchin EGF homolog (suEGF), human uromodulin (hURO), Drosophila crumbs (dCRU) and notch (dNOT), and xotch (xXOT), the Xenopus homolog of notch. Amino acid position within the full-length protein sequence is indicated at the right. EGF-like domains characterized by conserved cysteine residues are indicated above the sequences (EGF 1; EGF 2; EGF 3).

Table I. Pairwise Comparisons of % Identity and Alignment Score

|       | suEGF | dCRU | dNOT | xXOT |
|-------|-------|------|------|------|
| su63  | 47    | 47   | 43   | 40   |
| suEGF | 7.4   | 47   | 43   | 40   |
| dCRU  | 6.6   | 9.0  | 33   | 39   |
| dNOT  | 6.0   | 9.2  | 6.7  | 61   |
| xXOT  | 6.3   | 8.3  | 6.6  | 11.5 |

|       | hURO | dCRU | dNOT | xXOT |
|-------|------|------|------|------|
| su63  | 29   | 26   | 27   | 25   |
| hURO  | 10.0 | 33   | 37   | 37   |
| dCRU  | 5.4  | 6.5  | 33   | 37   |
| dNOT  | 5.1  | 7.8  | 10.7 | 61   |
| xXOT  | 4.0  | 8.1  | 13.8 | 25.2 |

Pairwise comparisons of percent amino acid identity (above the diagonals) and alignment score (below the diagonals) in SD units above random, as determined by the program ALIGN. The data in a and b correspond to the regions presented in Fig. 7 a and b. The matrix used for the alignment was PAM250; matrix bias was 6; gap penalty was 12.
63-kD protein is also a member of the EGF superfamily, sharing homology with the developmentally regulated *Drosophila* proteins crumbs (Tepass et al., 1990) and notch (Wharton et al., 1985), *Xenopus* protein xotch (Coffman et al., 1990), and the sea urchin EGF homolog (Hurst et al., 1987), as well as human uromodulin (Hession et al., 1987). The 63-kD protein shows no similarity to the two other cell membrane proteins of known sequence from sea urchin sperm: the speract receptor (Dangott et al., 1989) and guanylate cyclase (Singh et al., 1988).

GPI-anchored proteins are ubiquitous and have been implicated in a variety of physiological processes (Ferguson and Williams, 1988; Low and Saltiel, 1988; Thomas et al., 1990; Robinson, 1991). Of these, protein anchoring is the only clear function that can be assigned. GPI-anchored proteins reside exclusively in the apical membranes of polarized epithelial cells (Lisanti et al., 1988) and have been implicated in signal transduction in which inositol phosphate glycerol and myristylated diacylglycerol act as second messengers (Eardley and Koshland, 1991). Some effects of insulin (Saltiel et al., 1987), NGF (Chan et al., 1989; Mahanthappa and Patterson, 1992) and T cell activation (Gaulinon et al., 1988) are mediated by GPI-anchored proteins. Others have shown that protein tyrosine kinase activity is associated with GPI-anchored proteins on the surfaces of T cells (Stefanova et al., 1991; Thomas and Samelson, 1992).

GPI-anchoring of the 63-kD sea urchin sperm membrane protein, plus the fact that it contains 20 tyrosine residues are the most probable reasons why it is the most heavily labeled sperm protein when vectorial radioiodination is performed on living cells (Lopo and Vacquier, 1980a). The protein does not change location, nor detach from the cell, after the egg jelly induced AR. Significant homology to the developmentally regulated proteins crumbs, notch and xotch, suggests that the 63-kD may function in sperm cell differentiation. The cysteine-rich repeats characteristic of the EGF superfamily in proteins such as urokinase (Apella et al., 1987), laminin (Graf et al., 1987), coagulation factor IX (Rees et al., 1988) and notch (Rebay et al., 1991) are directly involved in protein–protein interactions underlying cell proliferation and differentiation.

Showing the most significant similarity with the sea urchin 63-kD protein is human uromodulin (Tamm-Horsfall protein). The alignment score of 10 SD U (Table I b) indicates that the probability of this homology occurring by chance is <1 in $10^{10}$. Additionally, like the 63-kD protein, uromodulin utilizes the modified polyadenylation site AATAAT (Hession et al., 1987; Fig. 1) and both proteins are GPI-anchored (Rindler et al., 1990; Fig. 6). The function of the 63-kD sea urchin sperm membrane protein remains unknown; however, it is worth noting that uromodulin, known for four decades to be the most abundant protein in human urine, remains unknown in terms of function (Kumar and Muchmore, 1990). The thick ascending limb of the loop of Henle in human kidney and the sea urchin sperm flagellum share the common attribute of being extremely active in ion flux. Gels of uromodulin can act as an electret, possessing a gross permanent dipole moment that allows the free passage of ions, but restricts the passage of water (Mattey and Naftalin, 1992). This observation is consistent with the observed high ionic and low water permeability of the thick ascending limb of the loop of Henle (Kumar and Muchmore, 1990). Similarly, in sea urchin sperm cells, high ion permeability and low water permeability across the flagellar membrane are crucial for the activation and maintenance of sperm motility (Clapper et al., 1985) and induction of the acrosome reaction (Schackmann and Shapiro, 1981; Shackmann et al., 1981; Christen et al., 1983; Lee, 1985). In relation to general phenomena of cellular homeostasis, protein homologous to the 63-kD protein of the sea urchin sperm flagellar membrane, and human uromodulin, might occur on the membranes of other cells that exist in high ionic strength media and also possess a relatively high rate of ionic flux.

We thank Dr. D. W. Smith for assistance with DNASYSTEM, Dr. R. F. Doolittle for design of the synthetic peptides, Dr. C. G. Glabe for peptide synthesis, Dr. A. V. Muchmore for antibody to uromodulin, Drs. N. Suzuki and L. J. Dangott for speract analogs, and Drs. E. H. Davidson and J. E. Minor for the original cDNA library.

L. M. Mendoza was supported by the Office of Graduate Studies, University of California, San Diego, and a National Science Foundation Minority Predoctoral Fellowship. D. Nishioka is supported by National Institutes of Health grant HD19054 and V. D. Vacquier National Institutes of Health grant HD12986.

Received for publication 3 February 1993 and in revised form 25 March 1993.

**References**

Apella, E., E. A. Robinson, S. J. Ullrich, M. P. Stoppelli, A. Corti, G. Cassani, and F. Blasi. 1987. The receptor-binding sequence of urokinase. *J. Biol. Chem.* 262:4437–4440.

Babcock, D. F., M. M. Bosma, D. E. Battaglia, and A. Darszon. 1992. Early development and differentiation. *The Journal of Cell Biology*, Volume 121, 1993 1296

Babcock, D. F., M. M. Bosma, D. E. Battaglia, and A. Darszon. 1992. Early cellular homeostasis, protein homologous to the 63-kD protein of the sea urchin sperm flagellar membrane, and human uromodulin, might occur on the membranes of other cells that exist in high ionic strength media and also possess a relatively high rate of ionic flux.

We thank Dr. D. W. Smith for assistance with DNASYSTEM, Dr. R. F. Doolittle for design of the synthetic peptides, Dr. C. G. Glabe for peptide synthesis, Dr. A. V. Muchmore for antibody to uromodulin, Drs. N. Suzuki and L. J. Dangott for speract analogs, and Drs. E. H. Davidson and J. E. Minor for the original cDNA library.

L. M. Mendoza was supported by the Office of Graduate Studies, University of California, San Diego, and a National Science Foundation Minority Predoctoral Fellowship. D. Nishioka is supported by National Institutes of Health grant HD19054 and V. D. Vacquier National Institutes of Health grant HD12986.

Received for publication 3 February 1993 and in revised form 25 March 1993.

**References**

Apella, E., E. A. Robinson, S. J. Ullrich, M. P. Stoppelli, A. Corti, G. Cassani, and F. Blasi. 1987. The receptor-binding sequence of urokinase. *J. Biol. Chem.* 262:4437–4440.

Babcock, D. F., M. M. Bosma, D. E. Battaglia, and A. Darszon. 1992. Early persistent activation of sperm K+ channels by the egg peptide speract. *Proc. Natl. Acad. Sci. USA.* 89:6001–6005.

Bloch, C. P., T. G. Wolfsberg, C. W. Turck, D. G. Myles, F. Primakoff, and J. M. White. 1992. A potential fusion peptide and an integrin ligand domain in a protein active in sperm-egg fusion. *Nature (London).* 356:248–252.
