Comparison of the Bindin Proteins of *Strongylocentrotus franciscanus, S. purpuratus,* and *Lytechinus variegatus:* Sequences involved in the Species Specificity of Fertilization[1]

Joseph E. Minor, David R. Fromson,† Roy J. Britten,* and Eric H. Davidson**

*Division of Biology, California Institute of Technology; and †Department of Biological Science, California State University, Fullerton

Downloaded from http://mbe.oxfordjournals.org/ at California Institute of Technology on December 18, 2015

Bindin is the sea urchin sperm acrosomal protein that is responsible for the species-specific adhesion of the sperm to the egg. Two new bindin cDNA sequences that contain the entire open reading frame for the bindin precursor are reported: one for *Strongylocentrotus franciscanus* and one for *Lytechinus variegatus.* Both contain inverted repetitive sequences in their 3′ untranslated regions, and the *S. franciscanus* cDNA contains an inverted repetitive sequence match between the 5′ untranslated region and the coding region. The middle third of the mature bindin sequence is highly conserved in all three species, and the flanking sequences share short repeated sequences that vary in number between the species. Cross-fertilization data are reported for the species *S. purpuratus, S. franciscanus, L. variegatus,* and *L. pictus.* A barrier to cross-fertilization exists between the sympatric *Strongylocentrotus* species, but there is no barrier between the allopatric *Lytechinus* species.

Introduction

Sea urchin fertilization involves a complex series of cell recognition events that insure fertilization specificity (Rossignol et al. 1984*a;* Minor et al. 1989). This specificity is necessary, since, like many marine invertebrates, sea urchins are broadcast spawners: gametes are released into open water where fertilization takes place. Sea urchins in general do not appear to go through premating isolating behaviors that would insure that the first egg a sperm would encounter would be one of the same species (Pennington 1985; Pearse et al. 1988). Therefore fertilization specificity is largely, if not entirely, dependent on the species specificity of gamete recognition. In the case of the species *Strongylocentrotus purpuratus* and *S. franciscanus,* the species specificity of fertilization has been shown to be due to the interaction between the sperm protein bindin and a glycoprotein receptor of the egg (Glabe and Vacquier 1977, 1978). Bindin is packaged as a large insoluble granule in the sperm acrosome. On contact with the egg, factors from the egg induce the sperm to undergo the acrosome reaction that exposes bindin. Bindin is not an integral membrane protein but remains associated with the sperm membrane and forms a contact with the receptor on the surface of the egg. In most cases, this contact must be species specific if efficient fertilization is to take place.

The bindin protein from *S. purpuratus* has been isolated and sequenced at the

1. Key words: fertilization, bindin, species specificity, inverted repeats, echinoderm.

Address for correspondence and reprints: Joseph E. Minor, Division of Biology, California Institute of Technology, Pasadena, California 91125.

Mol. Biol. Evol. 8(6):781–795. 1991.
© 1991 by The University of Chicago. All rights reserved.
0737-4038/91/0806-0004$02.00

protein level (Vacquier and Moy 1977, 1978; T. Sasagawa and K. Walsh, personal communication). A bindin cDNA for *S. purpuratus* was cloned using this sequence and was used to show that the *S. purpuratus* bindin message is 2.5 kb long and is a testis-specific transcript. The sequence of the clone revealed that bindin is initially synthesized as a 51-kDa precursor polypeptide which is cleaved to yield the 24-kDa acrosomal bindin from its C-terminus, plus a 27-kDa N-terminal protein of unknown function (Gao et al. 1986). The bindin gene is transcribed relatively late in spermatogenesis (Cameron et al. 1990; Nishioka et al. 1990) and is quickly translated and packaged into the acrosomal vesicle (Nishioka et al. 1990).

The four species of sea urchin used in the present study are *S. franciscanus, S. purpuratus, Lytechinus variegatus,* and *L. pictus.* The relationships between the species are summarized in table 1. The ranges of *S. franciscanus* and *S. purpuratus* overlap from British Columbia to Baja California on the Pacific coast of North America, and their breeding seasons overlap as well (Boolootian 1966). *Lytechinus variegatus* was chosen as an outgroup for comparison, since it is one of the most closely related species whose range does not overlap with that of the *Strongylocentrotus* species. *Lytechinus pictus* is from the Pacific coast, and its range overlaps with that of the *Strongylocentrotus* species. The two *Strongylocentrotus* species diverged 3.5–20 Mya, and the total single-copy divergence is 19%. The genus *Lytechinus* diverged from the genus *Strongylocentrotus* 30–40 Mya, and the total single-copy divergence is 31%. The sequences of the bindin cDNAs of *S. franciscanus* and *L. variegatus* are reported in the present paper. Also reported are data on cross-fertilization between *S. purpuratus, S. franciscanus, L. variegatus,* and *L. pictus.* The sequence data describe the changes that have taken place in bindin evolution, and the cross-fertilization data allow inferences to be made about how changes in the bindin protein have affected bindin function.

Material and Methods

Cross-Fertilizations

Gametes were collected by either electric shock or intracoelomic injection of 0.5 M KCl. The eggs were dejellied by washing in Millipore-filtered seawater (MPFSW) titrated to pH 5 by the addition of citric acid and then were returned to MPFSW. Then 500 eggs in 1 ml were placed in each of the wells of a Falcon 3047 cell culture plate. All work was done in an 18°C room. Sperm were diluted in MPFSW containing

Table 1
Relationships among Sea Urchin Species Used in Present Study

Species	Geographic Range[a]	Divergence Time[b] (Mya)	Divergence[c] (% scDNA)
Strongylocentrotus purpuratus	Pacific	...	4[d]
S. franciscanus	Pacific	3.5–20	19
Lytechinus pictus	Pacific	30–40	31
L. variegatus	Atlantic	30–40	ND

[a] Source: Boolootian (1966).

[b] Divergence times from *S. purpuratus* are based on a cladistic analysis of morphological data and on the fossil record from Smith (1988).

[c] Total single-copy DNA divergences are median divergence values reported by Hall et al. (1980) and M. S. Springer and R. J. Britten (unpublished observation). ND = not determined.

[d] Interindividual polymorphism within the species *S. purpuratus.*

Downloaded from http://bbx.oxfordjournals.org/ at California Institute of Technology on December 18, 2015

0.05% bovine serum albumin (BSA; Sigma Fraction V). A series of 1:1 dilutions were made, and aliquots of these dilutions were added to the eggs, producing the final dilution levels indicated (fig. 1). After 5 min, the eggs were fixed by the addition of an equal volume of 2% glutaraldehyde (Sigma) in MPFSW. The position of each species' eggs in the wells was randomized, and the plates were coded before scoring. Fertilization was scored by viewing the elevation of the fertilization envelope at 100 × in an inverted microscope. Dry-sperm concentrations were determined by counting in a Hausser chamber and by counting sperm fixed in the bottoms of the fertilization wells. The two values obtained were similar and were close to the value obtained from the yield of DNA extracted from dry sperm. The dry-sperm concentrations ($\times 10^{10}$ sperm/ml) were as follows: *Strongylocentrotus purpuratus,* 3.9; *S. franciscanus,* 9.5; *Lytechinus variegatus,* 4.6; and *L. pictus,* 4.0.

DNA and RNA Isolation, cDNA Synthesis, and Cloning

Sperm DNAs were isolated from *S. purpuratus* and *L. variegatus* by the method of Lee et al. (1984). *Strongylocentrotus franciscanus* sperm DNA was a gift from J. J. Lee and F. Calzone. Testis RNAs for *S. franciscanus* and *L. variegatus* were isolated according to a method described by Gao et al. (1986). *Strongylocentrotus purpuratus* testis RNA was a gift from B. Gao. PolyA$^+$ RNAs were purified by oligo-(dT) cellulose chromatography. *Strongylocentrotus franciscanus* cDNA was synthesized by the S1 hairpin method, while *L. variegatus* cDNA was synthesized using the RNAse H procedure (Amersham). *Eco*RI linkers were blunt-end ligated on the cDNA, and cDNAs were then size selected on a BioRad Biogel A-50m column and then cloned into λgt10 (Stratagene). The libraries were screened either with the total *S. purpuratus* cDNA (*S. franciscanus* library) or with the Sp5′ and Sp3′ fragments (defined below) of it (*L. variegatus* library). The hybridization was performed at 65°C in 5 × SET (SET = 0.15 M NaCl, 30 mM Tris-HCl pH8, 2 mM ethylenediaminetetraacetate), 5 × Denhardt's solution, and 25 μg sheared, denatured calf thymus DNA/ml. The final wash was in 1 × SET at 65°C. Multiple clones were isolated and subcloned into Bluescript vectors (Stratagene). Deletion subclones were made by the *Exo*III–mung bean nuclease deletion method, and one clone per species was sequenced completely on both strands. The sequenced clones were pBSf35 for *S. franciscanus* and pBLv22 for *L. variegatus.* The sequencing error is estimated to be 0.04% (sequences from each strand were obtained independently and differ by <2%).

DNA and RNA Gel Blot Analysis

Genomic DNAs were cleaved with a restriction enzyme (usually *Bgl*II), were separated by electrophoresis on agarose gels, and then were transferred to nitrocellulose. RNAs (both total and polyA$^+$) were separated by electrophoresis on formaldehyde gels and were transferred to nitrocellulose. Hybridizations were conducted as above, at either 65°C (final wash 1 × SET, 65°C) or 55°C (final wash 1 × SET, 55°C). The resulting filters were hybridized with several different DNA fragments taken from the *S. purpuratus* and *S. franciscanus* cDNAs. The *S. purpuratus* cDNA (see figure 1 in Gao et al. 1986) was divided into three fragments: Sp5′ (bases 1–637), SpMID (bases 638–1075), and Sp3′ (bases 1076–1873) by digestion with *Sal*I and *Eco*RI. The 3′ repetitive DNA fragment of the *S. franciscanus* cDNA (Sf3′) was isolated by digestion with *Bgl*II and *Eco*RI, yielding a 657-bp fragment (bases 1651–2307 in fig. 2).

Downloaded from http://mbe.oxfordjournals.org/ at California Institute of Technology on December 18, 2015

Results

Cross-Fertilizations

The degree of cross-fertilization between the four species (*Strongylocentrotus purpuratus, S. franciscanus, Lytechinus variegatus,* and *L. pictus*) was evaluated in the experiment described in figure 1. The general result is that barriers to cross-fertilization exist in most cases but can be overcome at very high sperm concentrations. The congeneric Pacific *Strongylocentrotus* species have overlapping geographical ranges, habitat preferences, and breeding seasons. Figure 1Sf shows that, even at the lowest sperm concentrations used, *S. franciscanus* sperm are effective at fertilizing their own eggs. Significant cross-fertilization with *S. purpuratus* eggs occurs only at 64-fold-higher sperm concentrations. As shown in figure 1Sp, *S. purpuratus* sperm are less effective at cross-fertilization, even at the highest concentrations of sperm used. As shown in figure 1Lv, *L. variegatus* sperm fertilize both species of *Lytechinus* well, and they also fertilize *S. franciscanus* equally well at higher sperm concentrations (the fertilization of *L. variegatus* eggs was inhibited at 18°C because their usual temperature is 25°C–30°C). Figure 1Lp shows that *L. pictus* sperm also fertilize both species of *Lytechinus* well but that they do not cross-fertilize *S. franciscanus* eggs as efficiently. Unlike the Pacific *Strongylocentrotus* species, the *Lytechinus* species (*L. variegatus* Atlantic and *L. pictus* Pacific) show no barrier to cross-fertilization.

In the sea urchin fertilization pathway there are three steps the failure of which could potentially result in the species specificity of fertilization. These steps are, in order of occurrence, movement of sperm to the egg, induction of the acrosome reaction, and adhesion of the sperm to the egg vitelline layer. It is this last step that is mediated by bindin. In a study of cross-fertilization between four sympatric species of sea urchins from Bermuda, Summers and Hylander (1975) concluded that in the majority (9/

Downloaded from http://biolbull.oxfordjournals.org/ at California Institute of Technology on December 18, 2015

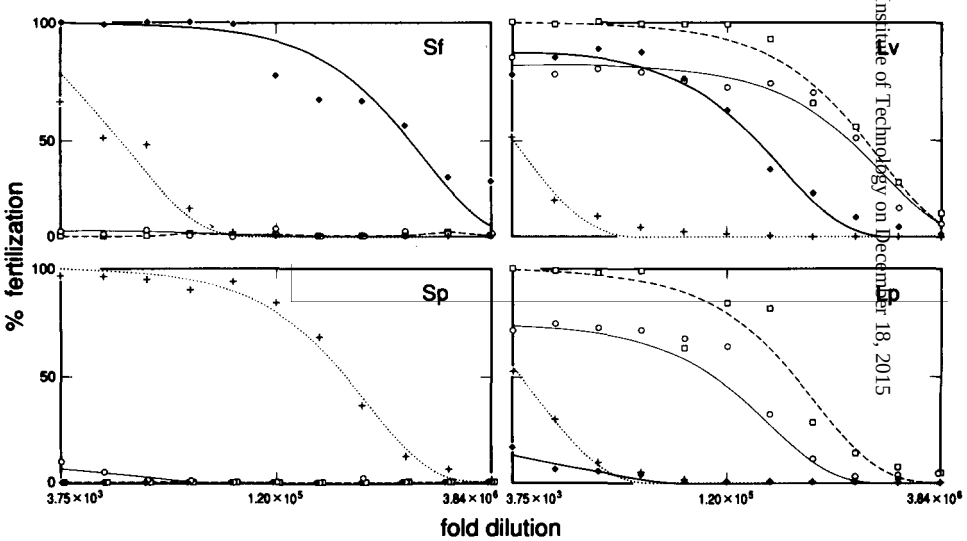


FIG. 1.—Cross-fertilization between *Strongylocentrotus franciscanus* (Sf), *S. purpuratus* (Sp), *Lytechinus variegatus* (Lv), and *L. pictus* (Lp). Each panel shows the result of fertilizing eggs of all four species (Sf, ♦——♦; Sp, +···+; Lv, ○——○; and Lp, □– – –□) with decreasing amounts of sperm from the species indicated in the upper-right corner. Eggs were dejellied and placed 500/1-ml well, and sperm were added at the final dilution indicated. *X*-axis is logarithmic; that is, each point has half the sperm concentration of the preceding point.

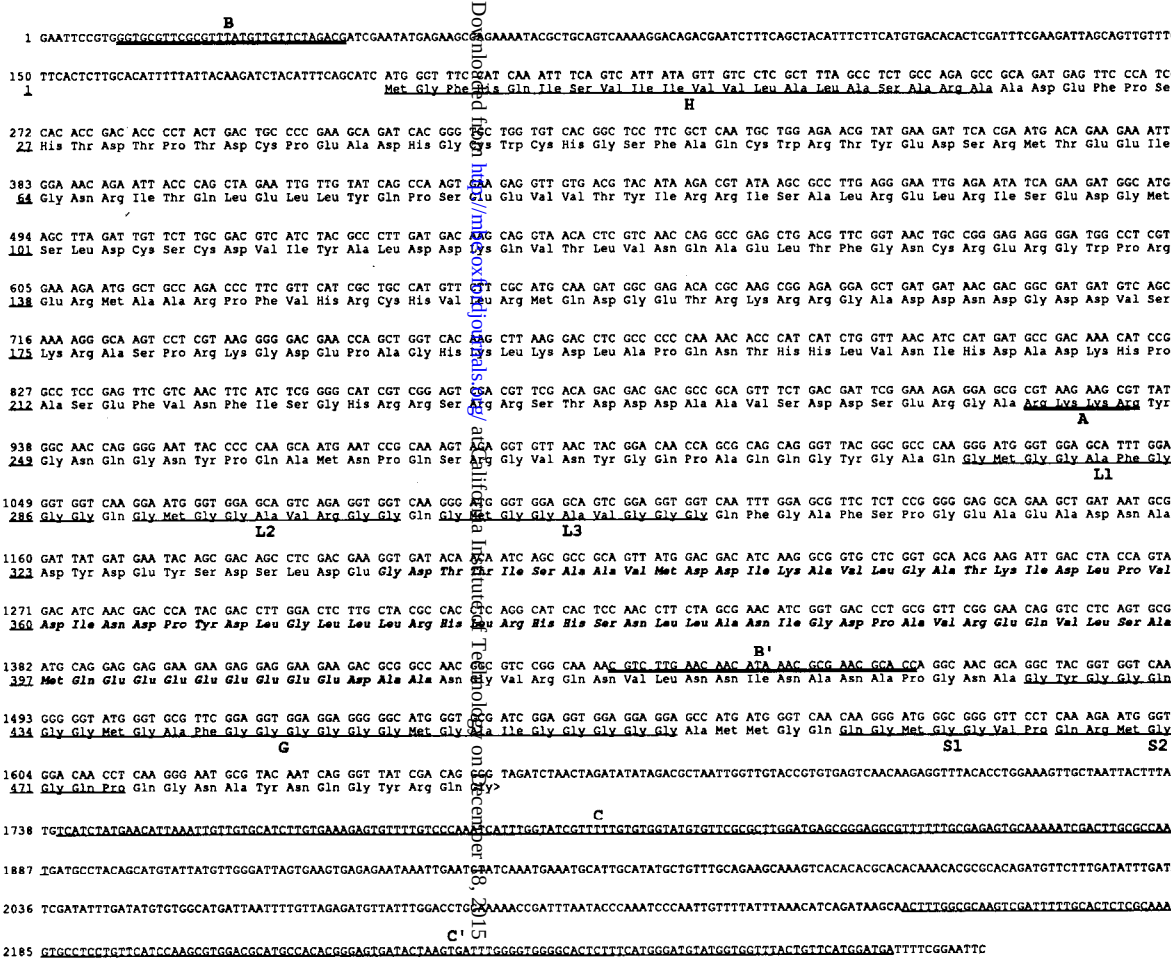
11) of cases the failure of cross-fertilization was due to a failure of sperm to adhere to the egg vitelline layer. However, there were two of 11 cases in which the failure of cross-fertilization was due to a failure of the acrosome reaction. It is also possible that species differences could exist in egg chemotactic molecules, which would result in species specificities in fertilization. An example of a sea urchin egg-derived chemotactic molecule is the *Arbacia* peptide resact (Ward et al. 1985). Species-specific egg chemotactic molecules are unlikely to be a concern for the species *S. purpuratus* and *S. franciscanus,* since the eggs of each species activate the motility of the sperm of the other species (Loeb 1915). The following considerations and observations show that neither chemotaxis nor acrosome reactions could have been responsible for the species-specific differences in efficiency of cross-fertilization seen in figure 1.

The possibility that the differences in fertilization could be due to a failure of chemotaxis, resulting in differing numbers of sperm reaching the eggs, was addressed by counting the number of sperm reaching the eggs. In the experiment described in figure 1, the eggs were fertilized, fixed, and scored in the same well to avoid the preferential loss of unfertilized eggs that occurs when eggs are transferred. After fertilization was scored, the number of sperm visible around the diameter of the unfertilized eggs was scored at 320 × [by the perimeter counting method of Kinsey et al. (1980)]. The sperm were counted around each of the four species of eggs at the sperm dilution that resulted in ~50% fertilization in the homotypic case. Since the eggs were not washed after fertilization, the sperm visible could be either sperm bound to the vitelline layer or sperm loosely "attached" in a nonspecific manner (see Summers and Hylander 1975; Wasserman 1990). There were no differences in the number of sperm per egg diameter, indicating no cross-species incompatibilities in the overall ability of sperm to locate the eggs.

The possibility that the species specificities of fertilization observed in figure 1 were due to failures of the acrosome reaction can be ruled out in 11 of the 12 interspecies crosses performed. There is no species specificity in the induction of the acrosome reaction either between the species *S. purpuratus* and *S. franciscanus* (A. Lopez and C. Glabe, personal communication) or between these species and *L. pictus* (B. Brandriff and V. Vacquier, personal communication). This eliminates the acrosome reaction as the cause of the species specificity of fertilization for the following crosses (in all crosses Sp = *S. purpuratus;* Sf = *S. franciscanus;* Lp = *L. pictus;* and Lv = *L. variegatus*): Sp♂ × Sf♀, Sp♂ × Lp♀, Sf♂ × Sp♀, Sf♂ × Lp♀, Lp♂ × Sp♀, and Lp♂ × Sf♀ (fig. 1Sf, Sp, and Lp). The acrosome reaction of *L. variegatus* is nonspecific (Summers and Hylander 1975) and, to a considerable degree, occurs spontaneously (SeGall and Lennarz 1979). This indicates that there are no acrosome-reaction compatibility problems for the crosses Lv♂ × Sp♀, Lv♂ × Sf♀, and Lv♂ × Lp♀ (fig. 1Lv). There is no acrosome-reaction incompatibility in the Lp♂ × Lv♀ cross, since this cross-fertilization proceeds efficiently (fig. 1Lp). *Lytechinus variegatus* egg jelly efficiently induces the acrosome reaction of *S. purpuratus* sperm (SeGall and Lennarz 1979), negating the possibility of a failure of the acrosome reaction in the Sp♂ × Lv♀ cross. Finally, the possibility of a failure of the acrosome reaction in the Sf♂ × Lv♀ cross cannot be excluded, but it seems unlikely in light of the general nonspecificity of the acrosome reaction in cross-fertilizations involving *S. franciscanus* sperm and *L. variegatus* eggs. Therefore, at least for 11 of the 12 cases tested, the only remaining cause for the species-specific inefficiencies in the cross-fertilizations shown in figure 1 is the failure of the sperm to bind productively with the egg.

In another experiment (data not shown), the barrier to cross-fertilization between

Downloaded from http://mbe.oxfordjournals.org/ at California Institute of Technology on December 18, 2015



Downloaded from http://mbe.oxfordjournals.org/ at California Institute of Technology on December 8, 2015

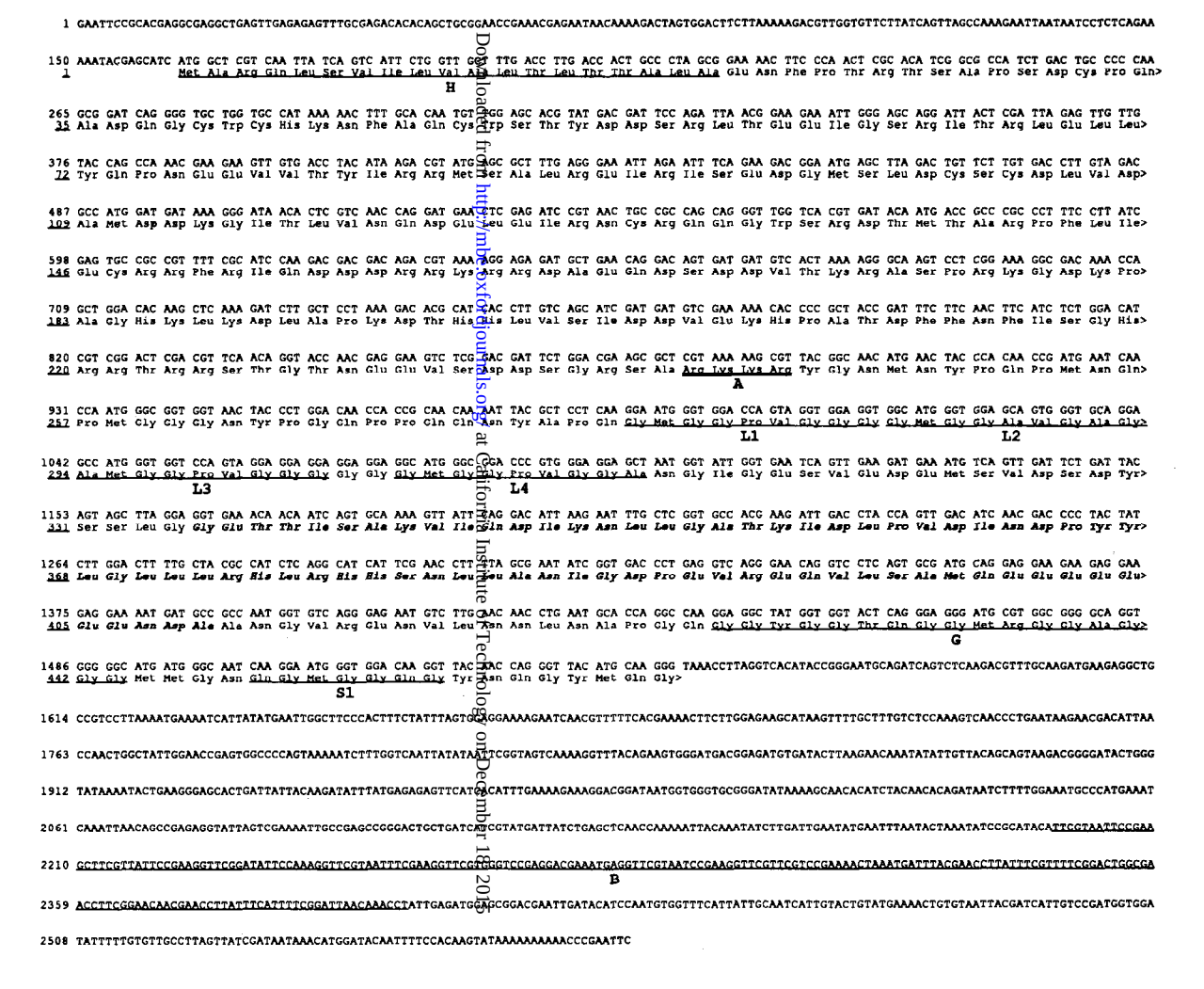
S. purpuratus and *S. franciscanus* was evaluated in detail by using eggs and sperm from different individuals, at limiting sperm concentration. Gametes were collected from four *S. franciscanus* females, five *S. franciscanus* males, five *S. purpuratus* females, and four *S. purpuratus* males. All possible interindividual fertilizations were performed at a sperm dilution that resulted in an average of 72% fertilization in the homologous cases (Sp♂ × Sp♀ or Sf♂ × Sf♀). Cross-fertilization (Sp♂ × Sf♀ or Sf♂ × Sp♀) rarely occurred and averaged <0.2%. These results show that under limiting sperm conditions, such as might be expected in the ocean (Pennington 1985), there is a very effective barrier to cross-fertilization between *S. purpuratus* and *S. franciscanus.*

Bindin Protein Sequences

Figure 2 displays the sequence of a 2.3-kb *S. franciscanus* bindin cDNA clone, and figure 3 displays that of a 2.6-kb *L. variegatus* cDNA clone. RNA gel blots (data not shown) show that the bindin message size is 3.3 kb for *S. franciscanus* and 3.0 kb for *L. variegatus.* Both clones contain the entire open reading frame for the bindin precursor, which in all three species is made as a large molecule approximately twice the size of the mature bindin. The predicted molecular weights for the bindin precursors are 52 kDa for *S. franciscanus* and 50 kDa for *L. variegatus.* The *S. franciscanus* precursor is known to be cleaved, since the acrosomal bindin has as its N-terminal residue the amino acid 248 of the precursor sequence (Vacquier and Moy 1978; also see fig. 2). On the basis of its sequence similarity to the other two species in the region surrounding the cleavage site, the *L. variegatus* precursor is assumed to be similarly cleaved, (figs. 3 and 4). The predicted molecular weights for the cleaved C-terminal acrosomal bindins are 24 kDa for *S. franciscanus* and 22 kDa for *L. variegatus.* The sequence preceding the cleavage site (RKKR; marked "A" in figs. 2 and 3) is conserved in all three species and constitutes a cleavage site for a protease that has trypsin-like specificity and is located in the Golgi apparatus (Strauss and Strauss 1985; Whealy et al. 1990). All three species of bindin contain repeats of two related protein sequences: GMGG(A/P)VGGG and QGMGG(P/Q) (P/H). These sequences are called the "long" (L) and "short" (S) repeats, and all the indicated sequences (figs. 2, 3, and 5) have, at most, one difference from the consensus sequence. The middle third of the bindin molecule is highly conserved in all three species (italicized sequences in figs. 2 and 3).

FIG. 2.—Sequence and translation of *Strongylocentrotus franciscanus* bindin cDNA. The sequence is 2,307 bp in length (including the *Eco*RI cloning linkers). The sequences are numbered to the left, and the numbers for the protein sequences are underlined. The open reading frame extends from base 194 to base 1648. A hydrophobic leader sequence (underlined and marked "H") extends from amino acid 1 to amino acid 20, where it ends with a putative leader cleavage site (von Heijne 1983, 1984). The double-underlined sequences marked "A" denote a signal sequence for protease cleavage that precedes the cleavage between amino acids 247 and 248 that separates bindin from the N-terminal protein. L1, L2, and L3 (underlined) denote repeats of the long consensus sequence GMGG(A/P)VGGG. S1 and S2 (underlined) denote repeats of the short consensus sequence QGMGG(P/Q)(P/H). A highly conserved block of sequence (italicized) extends for 76 amino acids, from 334 to 409. "G" denotes a stretch of glycine residues. The double-underlined DNA sequences (B and B′) denote an inverted repeat (29-nt/30-nt match) that covers part of the coding region. C and C′ (underlined) denote a 149-nt inverted repeat (with stretches of perfect match up to 35 nt) that is in the 3′ untranslated region. This sequence has been submitted to GenBank under accession number M59490.

Downloaded from http://mbe.oxfordjournals.org/ at California Institute of Technology on December 18, 2015



Downloaded from http://mbe.oxfordjournals.org/ at California Institute Technology on December 16, 2015

Inverted and Repetitive Sequences on Bindin mRNAs

DNA gel blots using the SpMID fragment indicate that bindin is a single-copy gene in all three species (data not shown). Figure 4A shows that the *S. franciscanus* cDNA contains a repetitive sequence that is not present on the *S. purpuratus* cDNA. This repetitive sequence is present in the genomes of *S. franciscanus* (fig. 4, lane 1) and *S. purpuratus* (data not shown) but is absent from the genome of *L. variegatus* (fig. 4B, lanes 2 and 3). This repetitive sequence was shown to be limited to the 3′ untranslated region of the *S. franciscanus* mRNA, by using a probe (Sf3′) that contains only 3′ untranslated sequences. When the Sf3′ probe was used on gel blots of testis RNA, it was shown that this repetitive sequence is present only on the bindin mRNA of *S. franciscanus,* not on either any other testis RNA of *S. franciscanus* or any testis RNA (including bindin) of the other two species. The sequence of the *S. franciscanus* bindin cDNA reveals that this repetitive element (C and C′ in fig. 2) consists of a large inverted repeat. The size of this repeat accounts for most of the size difference between the bindin mRNAs of the two *Strongylocentrotus* species. Sequence analysis of the *L. variegatus* cDNA reveals a different 3′ untranslated repetitive sequence (B in fig. 3). The simplest explanation of these data is that a novel repetitive element present in the *Strongylocentrotus* lineage spread through the genomes after the divergence of the two species. During this process this repetitive element landed in the *S. franciscanus* bindin transcription unit but is not found in the *S. purpuratus* bindin, though, as noted above, it is present elsewhere in the *S. purpuratus* genome. The *L. variegatus* bindin transcription unit would then have independently acquired a different 3′ repetitive element.

The *S. franciscanus* cDNA contains a second inverted DNA sequence (B and B′ in fig. 2) that consists of a 29-of-30-nt match between a sequence in the 5′ untranslated region and its complement in the coding region. This "antisense pair," as well as the other repetitive elements, could affect both transcript stability and translation control of bindin. The functional significance of all of these repeats, as well as the possibility of additional repetitive elements on the uncloned portions of all three bindin mRNAs, remains uninvestigated.

Phylogenetic Distribution of Bindin

A search was made for the presence of bindin DNA sequences in species other than sea urchins by using fragments of the *S. purpuratus* cDNA (Sp5′, SpMID, and Sp3′) at the 55°C criterion to probe "zoo blots" (blots containing the *Bgl*II-digested genomic DNAs of three sea urchin species (*S. purpuratus, S. franciscanus,* and *L. variegatus*), a starfish (*Pisaster ochraceus*), a sea cucumber (*Thyone briareus*), *Xenopus*

FIG. 3.—Sequence and translation of *Lytechinus variegatus* bindin cDNA. The sequence is 2,584 bp in length (including the *Eco*RI cloning linkers). The open reading frame extends from base 163 to base 1548. A hydrophobic leader sequence (underlined and marked "H") extends from amino acid 1 to amino acid 19, where it ends with a putative leader cleavage site. The double-underlined sequences marked "A" denote a signal sequence for protease cleavage that precedes the putative cleavage between amino acids 242 and 243 that separates bindin from the N-terminal protein. L1–L4 (underlined) denote repeats of the long consensus sequence. S1 (underlined) denotes a short consensus-sequence site. A highly conserved block of sequence (italicized) extends for 76 amino acids, from 335 to 410. "G" denotes a stretch of glycine residues. The underlined sequence "B" from base 2196 to 2404 denotes an area in the 3′ untranslated region and consisting of multiple direct and inverted repeats of the sequence TTCG. This sequence has been submitted to GenBank under accession number M59489.

Downloaded from http://mbe.oxfordjournals.org/ at California Institute of Technology on December 18, 2015

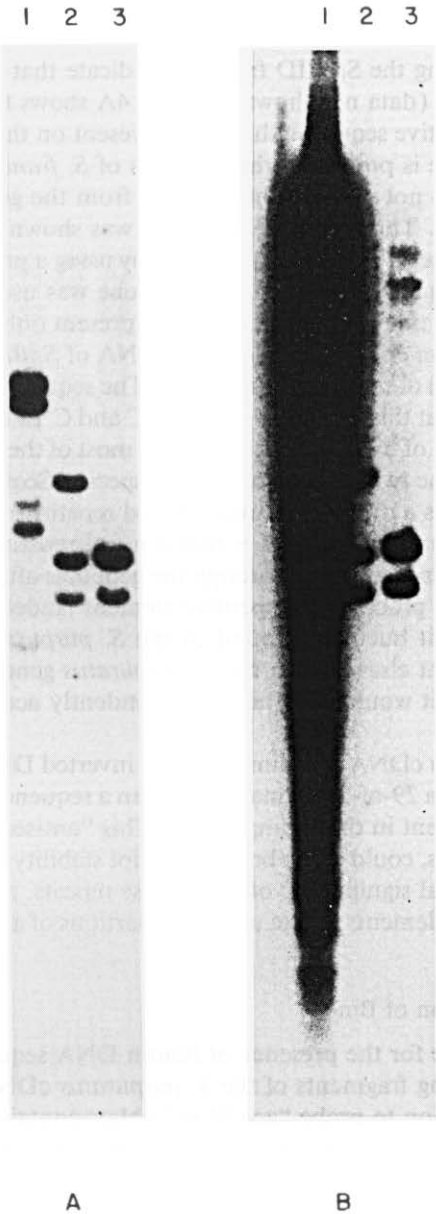


FIG. 4.—Genomic DNA blots probed with bindin cDNAs. Sperm genomic DNAs from a *Strongylocentrotus franciscanus* individual (lanes 1) and two *Lytechinus variegatus* individuals (lanes 2 and lanes 3) were cut with *Bgl*II, separated on an agarose gel, and transferred to nitrocellulose. The filter was then hybridized to labeled cDNA from the entire *S. purpuratus* bindin cDNA (A) or the entire *S. franciscanus* bindin cDNA (B). The bindin cDNAs are single-copy sequences in *Lytechinus*, and the *S. franciscanus* cDNA contains a sequence that is repetitive in the genome of *S. franciscanus*.

laevis, Mus domesticus, Homo sapiens, Aplysia californica, Mytilus edulis, Loligo pealii, Drosophila simulans, Caenorhabditis elegans, Cerebratulus lacteus, Tetrahymena thermophila, Macrocystis integrifolia, and *Saccharomyces cerevisiae.* No reliable

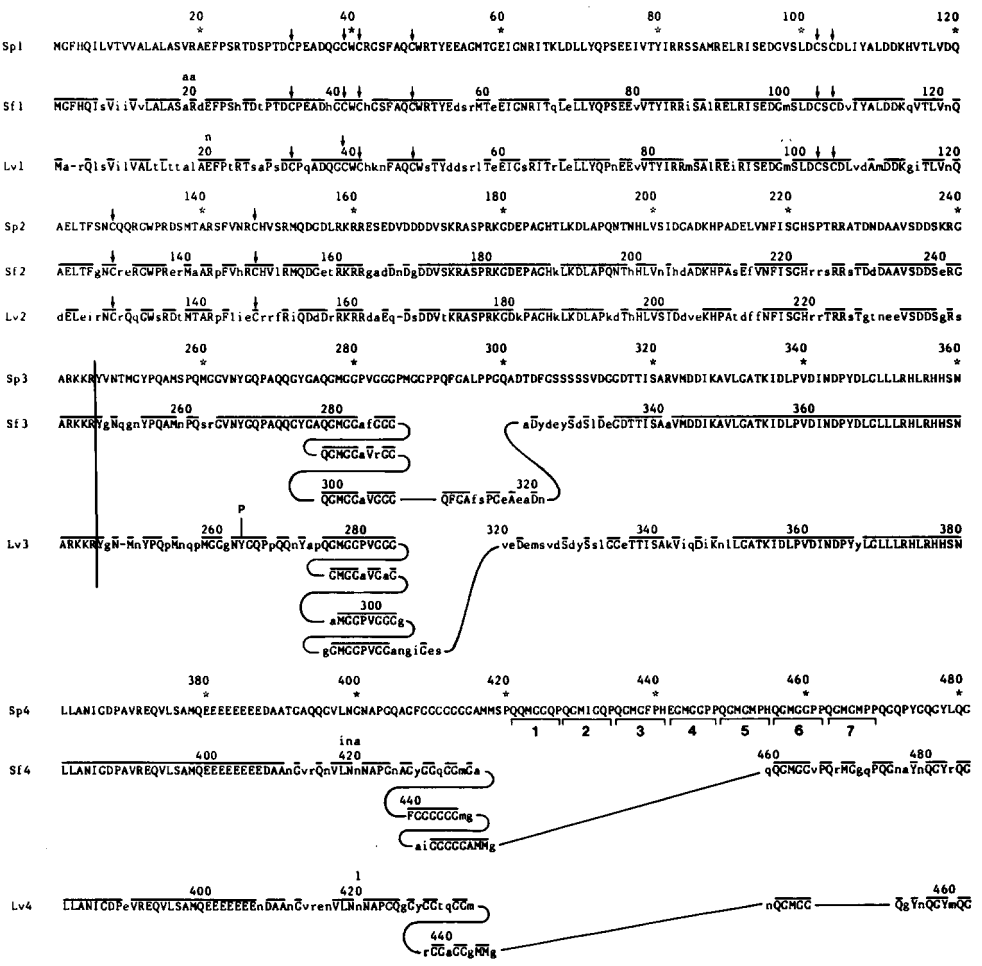


FIG. 5.—Comparison of entire bindin precursor sequences of *Strongylocentrotus purpuratus* (Sp), *S. franciscanus* (Sf), and *Lytechinus variegatus* (Lv). The sequences are aligned relative to *S. purpuratus;* amino acid identities are given in overlined capital letters, and mismatches are given in lower case letters. Deletions in sequence relative to *S. purpuratus* are denoted by dashes (-), while insertions are shown, right justified, above the amino acid that they follow. Large repeats in sequence are shown underneath each other, and their order in the sequence is shown connected by lines. The vertical line that follows amino acid 245 in Sp marks the cleavage site that frees the mature acrosomal bindin from the precursor. Numbers 1–7 appearing beneath the *S. purpuratus* sequence after amino acid 420 denote repeats of the short consensus sequence.

signal was found outside the sea urchin species, a result consistent with earlier antibody studies (Moy and Vacquier 1979). Bindin (and bindin-precursor) sequences seem to be limited to the echinoids.

Discussion

The derived sequences from the entire open reading frames of three species of bindin are aligned in figure 5. There is a dramatic difference in the pattern of sequence evolution when the N-terminal protein sequences (amino acids 1–245 of *Strongylocentrotus purpuratus*) are compared with the mature bindin sequences (amino acids

Downloaded from http://mbe.oxfordjournals.org/ at California Institute of Technology on December 18, 2015

246–481 of *S. purpuratus*). In the N-terminal protein sequences, there are few sequence rearrangements (insertions, deletions, or repeated sequences) but many point changes, while in the mature bindin there are extensive sequence rearrangements but comparatively fewer point changes. This is a variant of the "mosaic evolution" pattern seen in the *Drosophila* sgs genes (Martin and Meyerowitz 1988), where there is a 5–10-fold change in the frequency of nucleotide substitutions on either side of a boundary that is <50 nt long. In the case of bindin and N-terminal sequences, the differences in the pattern of sequence evolution probably reflect the effects of selection for different functions on different parts of the protein. While the function of the N-terminal protein sequences remains unknown, it has been speculated, on the basis of sequence content, that the N-terminal protein might play a role in packaging the insoluble mature bindin for transport to the acrosome (Gao et al. 1986). The fact that, despite multiple point changes, the three species of bindin precursor all conserve their overall size and structure, including the exact relative positioning of eight cysteines (fig. 5, vertical arrows), tends to support this notion.

The extensive rearrangement of the mature bindin sequences directly contrasts with the above pattern. The middle third of the molecule is highly conserved (there is a stretch of 68 amino acids that is perfectly conserved within *Strongylocentrotus*). This conserved region [amino acids 314–390 of *S. purpuratus* (fig. 5), diagrammed as a black box in fig. 6] corresponds to the region that in *S. purpuratus* is thought to bind sulfated fucans (DeAngelis and Glabe 1988, 1990) and to possibly interact with the sperm acrosomal membrane (Kennedy et al. 1989). There are three areas of sequence rearrangement between the three species. The first involves a polyglycine stretch (G in figs. 2, 3, and 6) that is longer in *S. franciscanus* and *Lytechinus variegatus* than in *S. purpuratus*. The second and third areas involve differences in the number and position of two related repeated protein sequences: GMGG(A/P)VGGG (L in fig. 6) and QGMGG(P/Q) (P/H) (S in fig. 6). The difference in the number of L repeats between *S. purpuratus* and *S. franciscanus* was first noticed in N-terminal protein sequences of bindin determined by Vacquier and Moy (1978). Around the conserved repetitive regions are short areas of species-unique sequence (e.g., amino acids 288–314 of Sp or amino acids 312–333 of Sf; fig. 4). Presumably some of these

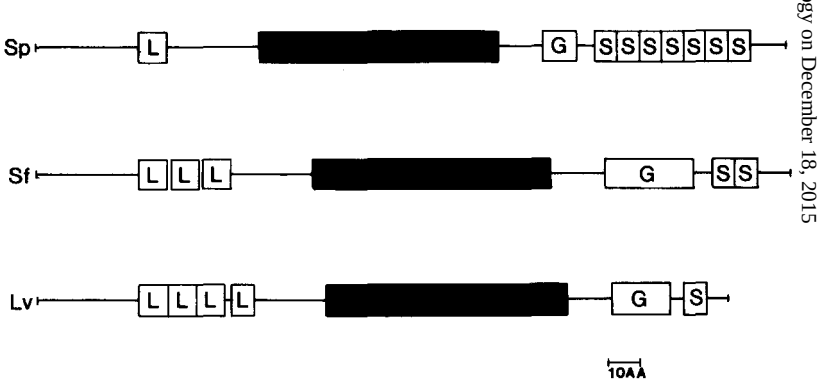


FIG. 6.—Diagram of protein sequence relationships between the bindins of *Strongylocentrotus purpuratus* (Sp), *S. franciscanus* (Sf), and *Lytechinus variegatus* (Lv). The major block of conserved sequence is shown as a black box at the core of the three bindins. Tracts of polyglycine are denoted by a "G." "L" denotes repeats of the long consensus sequence GMGG(A/P)VGGG, and "S" denotes repeats of the short sequence QGMGG(P/Q)(P/H).

Downloaded from http://mbe.oxfordjournals.org/ at California Institute of Technology on December 18, 2015

sequence differences are responsible for the fertilization-specificity differences seen between the species.

The cDNA sequence has recently been reported for the bindin of a fourth species of sea urchin, *Arbacia punctulata* (Glabe and Clark 1991). *Arbacia* is a distant relative of the sea urchins in the present study, having diverged from them ~200 Mya (Smith 1984). Despite this long divergence time, the *Arbacia* bindin shares several features with the other bindins. The *Arbacia* bindin is also synthesized as a larger precursor, and the N-terminal sequences conserve the eight cysteines found in the *Strongylocentrotus* and *Lytechinus* bindins. It also conserves a large section of the core sequences conserved among the species in the present study. The *Arbacia* bindin differs from the bindins reported here in that it contains an extra hydrophobic region, and it does not conserve the L and S sequences.

The uniqueness of bindin does not allow predictive structural comparisons with other recognition proteins. Although it mediates cell-cell recognition and possibly membrane fusion (Glabe 1985*a*, 1985*b*), bindin is not an integral membrane protein. Bindin is insoluble in seawater, and in electron micrographs it is visible as a "blob" of protein both in the sperm acrosomal vesicle (Nishioka et al. 1990) and between the sperm and egg at fertilization (Moy and Vacquier 1979). In addition, it binds in a strong field of contact (capable of holding a moving sperm) that involves many bindin monomers. The *S. purpuratus* bindin receptor is a high-molecular-weight ($>10^7$-Da) complex that requires both protein and carbohydrate for its specificity (Rossignol et al. 1984*b*; Foltz and Lennarz 1990). The bindin-receptor interaction takes place in seawater, which has a higher ionic strength than that in which recognition proteins with solved structures operate. It is not surprising, then, that bindin is unrelated, in sequence, to other proteins: bindin and bindin precursor sequences were searched against both the National Biomedical Research Foundation's Protein Identification Resource Database (releases 22.0 and 40.0, September 1989) and GenBank (release 63, March 1990), and no significant matches were found.

The differences in the bindin proteins could result in changes in the species specificity of fertilization if some of the altered sequences encoded species-specific domains of interaction between bindin and its receptor. This idea has been tested in our laboratory by assaying the ability of bindin-derived peptides to inhibit fertilization. We have found that fertilization is inhibited in a species specific manner (authors' unpublished observation) by a peptide derived from a unique region of the *S. franciscanus* bindin, a result supporting the view that the species specificity of the bindin-receptor interaction is due to altered binding domains. It will be interesting not only to continue these structural studies but to study the population distribution of bindin alleles in groups of species and subspecies that may be currently undergoing speciation (e.g., the *Echinometra* species complex; Lessios and Cunningham 1990; Palumbi and Metz 1991).

Sequence Availability

The sequences shown in figures 2 and 3 have been deposited in GenBank under accession numbers M59490 and M59489, respectively.

Acknowledgments

We thank Roger Anderson, Dr. D. Nishioka, and the two reviewers for helpful comments on the manuscript. We also thank Drs. F. Calzone, S. Fain, H. Hwu, S.

Downloaded from http://mbe.oxfordjournals.org/ at California Institute of Technology on December 18, 2015

Kim, D. Livant, F. Preugschat, and M. J. Smith and S. Westaway for genomic DNAs used in the zoo blots. This work was supported by NSF grant DCB-8813413.

LITERATURE CITED

BOOLOOTIAN, R. A. 1966. Reproductive physiology. Pp. 561–614 *in* Physiology of echinodermata. Interscience, New York.

CAMERON, R. A., J. E. MINOR, D. NISHIOKA, R. J. BRITTEN, and E. H. DAVIDSON. 1990. Locale and level of bindin mRNA in maturing testis of the sea urchin *Strongylocentrotus purpuratus*. Dev. Biol. **142**:44–49.

DEANGELIS, P. L., and C. G. GLABE. 1988. Role of basic amino acids in the interaction of bindin with sulfated fucans. Biochemistry **27**:8189–8194.

———. 1990. Zinc specifically stimulates the selective binding of a peptide analog of bindin to sulfated fucans. Peptide Res. **3**:62–68.

FOLTZ, K. R., and W. J. LENNARZ. 1990. Purification and characterization of an extracellular fragment of the sea urchin egg receptor for sperm. J. Cell Biol. **111**:2951–2959.

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.

GLABE, C. G. 1985*a*. Interaction of the sperm adhesive protein, bindin, with phospholipid vesicles. I. Specific association of bindin with gel-phase phospholipid vesicles. J. Cell Biol. **100**:794–799.

———. 1985*b*. Interaction of the sperm adhesive protein, bindin, with phospholipid vesicles. II. Bindin induces the fusion of mixed-phase vesicles that contain phosphatidylcholine and phosphatidylserine in vitro. J. Cell Biol. **100**:800–806.

GLABE, C. G., and D. CLARK. 1991. The sequence of the *Arbacia punctulata* bindin cDNA and implications for the structural basis of species-specific sperm adhesion and fertilization. Dev. Biol. **143**:282–288.

GLABE, C. G., and V. D. VACQUIER. 1977. Species specific agglutination of eggs by bindin isolated from sea urchin sperm. Nature **267**:822–824.

———. 1978. Egg surface glycoprotein receptor for sea urchin sperm bindin. Proc. Natl. Acad. Sci. USA **75**:881–885.

HALL, T. J., J. W. GRULA, E. H. DAVIDSON, and R. J. BRITTEN. 1980. Evolution of sea urchin non-repetitive DNA. J. Mol. Evol. **16**:95–110.

KENNEDY, L., P. L. DEANGELIS, and C. G. GLABE. 1989. Analysis of the membrane-interacting domain of the sea urchin sperm adhesive protein bindin. Biochemistry **28**:9153–9158.

KINSEY, W. H., J. A. RUBIN, and W. J. LENNARZ. 1980. Studies on the specificity of sperm binding in echinoderm fertilization. Dev. Biol. **74**:245–250.

LEE, J. J., R. J. SHOTT, S. J. ROSE III, T. L. THOMAS, R. J. BRITTEN, and E. H. DAVIDSON. 1984. Sea urchin actin gene subtypes: gene number, linkage, and evolution. J. Mol. Biol. **172**:149–176.

LESSIOS, H. A., and C. W. CUNNINGHAM. 1990. Gametic incompatibility between species of the sea urchin *Echinometra* on the two sides of the isthmus of Panama. Evolution **44**:933–941.

LOEB, J. 1915. On the nature of the conditions which determine or prevent the entrance of the spermatozoan into the egg. Am. Nat. **49**:257–285.

MARTIN, C. H., and E. M. MEYEROWITZ. 1988. Mosaic evolution in the *Drosophila* genome. Bioessays **9**:65–69.

MINOR, J. E., B. GAO, and E. H. DAVIDSON. 1989. The molecular biology of bindin. Pp. 73–88 *in* H. SCHATTEN and G. SCHATTEN, eds. The molecular biology of fertilization. Academic Press, San Diego.

Downloaded from http://dbiox.oxfordjournals.org/ at California Institute of Technology on December 1, 2015

MOY, G. W., and V. D. VACQUIER. 1979. Immunoperoxidase localization of bindin during the adhesion of sperm to sea urchin eggs. Top. Dev. Biol. **13**:31–44.

NISHIOKA, D., R. A. WARD, D. POCCIA, C. KOSTACOS, and J. E. MINOR. 1990. Localization of bindin expression during sea urchin spermatogenesis. Mol. Reprod. Dev. **27**:181–190.

PALUMBI, S. R., and E. C. METZ. 1991. Strong reproductive isolation between closely related tropical sea urchins (genus *Echinometra*). Mol. Biol. Evol. **8**:227–239.

PEARSE, J. S., D. J. MCCLARY, M. A. SEWELL, W. C. AUSTIN, A. PEREZ-RUZAFA, and M. BYRNE. 1988. Simultaneous spawning of six species of echinoderms in Barkley Sound, British Columbia. Invertebrate Reprod. Dev. **14**:279–288.

PENNINGTON, J. T. 1985. The ecology of fertilization of echinoid eggs: the consequences of sperm dilution, adult aggregation, and synchronous spawning. Biol. Bull. **169**:417–430.

ROSSIGNOL, D. P., G. L. DECKER, and W. J. LENNARZ. 1984*a*. Cell-cell interactions and membrane fusion during fertilization in sea urchins. Pp. 5–26 *in* R. F. BEERS and E. G. BASSET, eds. Cell fusion: gene transfer and transformation. Raven, New York.

ROSSIGNOL, D. P., B. J. EARLES, G. L. DECKER, and W. J. LENNARZ. 1984*b*. Characterization of the sperm receptor on the surface of the eggs of *Strongylocentrotus purpuratus*. Dev. Biol. **104**:308–321.

SEGALL, G. K., and W. J. LENNARZ. 1979. Chemical characterization of the component of the jelly coat from sea urchins responsible for induction of the acrosome reaction. Dev. Biol. **71**:33–48.

SMITH, A. B. 1984. Echinoid paleobiology. Vol. **1** *in* C. T. SCROTON and C. P. HUGHES, eds. Special topics in paleontology. Allen & Unwin, London.

———. 1988. Phylogenetic relationship, divergence times, and rates of molecular evolution for camarodont sea urchins. Mol. Biol. Evol. **5**:345–365.

STRAUSS, E. G., and J. H. STRAUSS. 1985. Assembly of enveloped animal viruses. Pp. 205–234 *in* S. CASJENS, ed. Virus structure and assembly. Jones & Bartlett, Portola Valley, Calif.

SUMMERS, R. G., and B. L. HYLANDER. 1975. Species-specificity of acrosome reaction and primary gamete binding in echinoids. Exp. Cell Res. **96**:63–68.

VACQUIER, V. D., 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.

———. 1978. Macromolecules mediating sperm-egg recognition and adhesion during sea urchin fertilization. Pp. 379–389 *in* E. R. DIRKSEN, ed. Cell reproduction: in honor of Daniel Mazia. Academic Press, New York.

VON HEIJNE, G. 1983. Patterns of aminoacids near signal-sequence cleavage sites. Eur. J. Biochem. **133**:17–21.

———. 1984. How signal sequences maintain cleavage specificity. J. Mol. Biol. **173**:243–251.

WARD, G. E., C. J. BROKAW, D. L. GARBERS, and V. D. VACQUIER. 1985. Chemotaxis of *Arbacia punctulata* spermatozoa to resact, a peptide from the egg jelly layer. J. Cell Biol. **101**:2324–2329.

WASSERMAN, P. M. 1990. Profile of a mammalian sperm receptor. Development **108**:1–17.

WHEALY, M. E., A. K. ROBBINS, and L. W. ENQUIST. 1990. The export pathway of the pseudorabies virus gB homolog gII involves oligomer formation in the endoplasmic reticulum and protease processing in the Golgi apparatus. J. Virol. **64**:1946–1955.

RICHARD K. KOEHN, reviewing editor

Received January 23, 1991; revision received May 7, 1991

Accepted May 7, 1991

Downloaded from http://mbe.oxfordjournals.org/ at California Institute of Technology on December 18, 2015