The Structure and Expression of the Gene Coding for Bindin, a Species Specific Sea Urchin Sperm Protein

Thesis by

Boning Gao

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1986

(Submitted April 11, 1986)

Acknowledgements

My years at Caltech have been most scientifically rewarding and personally enriching. This was made possible by many people at Caltech. I particularly want to thank my advisor, Eric Davidson, whose guidance and patience have nurtured my growth as a researcher. I thank all of the members of my Committee for their valuable suggestions on my thesis. I also wish to thank all of the members in the Davidson group, past and present, especially Laurie Klein and Jamey Ferrier for their work on the bindin project, without them the work would have been much slower; Barbara Hough-Evans, Andy Cameron, and Henry Sucov for their detailed criticism of my thesis; Costas Flytzanis, Jamie Lee, and Frank Calzone for their "trouble shooting" suggestions on solving technical problems; Jane Rigg and Barbara Hough-Evans for their friendship and constant help from the first day of my life at Caltech. I must thank Joe Minor and Charles Glabe for their stimulating discussions which are the basis for the last chapter of my thesis. I am indebted to Dr. Vacquier, whose excellent work on bindin protein inspired me to initiate this thesis work and whose generous help made this work possible. I would like to thank my former advisor, Shen Chen, for his encouragement over the years. I gratefully acknowledge the secretaries in the Biology Office for their time and effort in keeping up with my frequent changes and corrections on this thesis. Finally, I thank my family and my husband for their patience and unfailing source of support. I was supported by a fellowship from the California Foundation for Biochemical Research. I received funds for the preparation of this thesis from the Jean Weigle Memorial Fund.

Abstract

Bindin is a major protein of the sea urchin sperm acrosome granule which mediates the species-specific adhesion and binding of sperm to the egg. Bindin protein has been purified from the sperm of the sea urchin *Strongylocentrotus purpuratus* (*S. purpuratus*) and the protein has been partially sequenced. The work presented in this thesis is the isolation and the sequence analysis of bindin cDNA and gene, and the study of the expression of the bindin gene.

A λgt10 cDNA library was constructed from *S. purpuratus* testes poly(A)$^+$ RNA. The library which was screened with a synthetic DNA probe prepared according to the known protein sequence yielded clones representing bindin cDNA. One of these clones containing an 1873 base pair (bp) insert was sequenced and found to code for a bindin precursor (prebindin) which is twice as large as the mature bindin protein. Upon immunoprecipitation with bindin antibody, the testes poly(A)$^+$ RNA *in vitro* translation product yields a larger precursor. The bindin cDNA was used to study the tissue specificity of expression. The results show that bindin is a sperm specific protein—its messenger RNA is not detected in the eggs, ovaries, early embryos, coelomocytes, tubefeet, lantern tissue or intestine that we tested.

Sperm DNA isolated from several individuals was probed with bindin cDNA and reveals that there is one bindin gene per haploid genome. Haploid female coelomocyte DNA also possesses one bindin gene.

Bindin cDNA was used to screen an EcoRI partially digested sea urchin *S. purpuratus* DNA charon 4A library. A clone containing a 14 kilobase (Kb) insert hybridized to the 3' end of the bindin cDNA. It also has overlapping restriction enzyme recognition sites with the 3' end of the bindin cDNA. There is an intron in

the genomic clone upstream of this overlapping region since it did not hybridize with bindin cDNA. A λgt10 mini-genomic library was made and a 1.3 Kb genomic DNA clone which hybridized with bindin cDNA has been characterized and partially sequenced. It contains a 219 bp exon in which the 5' end lies 2 bp upstream of the AUG translation initiation codon. This exon is flanked by introns on either side. Thus, there are at least three introns in the bindin gene.

Table of Contents

INTRODUCTION

Introduction

Fertilization is the inception of a new generation. It is the fusion of two gametic cells, male and female, followed by the joining together of the nuclei of the two gametes. Fertilization is a complicated process which involves cell-cell recognition, cell attachment, cell fusion, pronuclei fusion, a series of intricate ionic movements, the polymerization of actin, exocytosis, endocytosis and numerous other physiological changes in the two gametes and the zygote. Nature has invented many kinds of devices to carry out this process to facilitate the interaction of homologous eggs and sperm and to prevent the heterologous fertilization which in most cases will lead to failure in early development.

There are two basic patterns of fertilization in the animal kingdom. In the case of external fertilization, both the egg and the sperm are discharged into the surrounding medium. The molecules on the surface of eggs and sperm are responsible for specificity in the external fertilization. The sperm is introduced into the genital ducts of the female in the case of internal fertilization. In this case, the behavioral barriers and the compatibility of the genital organs preserve a specificity of fertilization, and the individuality of the characteristics of the gamete surface molecules may be less significant.

A successful fertilization consists of the following events: 1) the encounter of the two gametes, 2) the acrosome reaction of the sperm, 3) the binding of the sperm to the egg, 4) the fusion of the two plasma membranes, and thereafter, 5) the fusion of the two pronuclei. These sequential events were initially investigated by Colwin and Colwin (1-4) based on the election microscopic study of the fertilization in Hydroides (Annelida) and Saccoglossus (Hemichordata). The same pattern has been observed in many animals including echinoderms (5-8), mollusks (7,9-11), echiuroids (12), and mammals (13-20).

Fertilization in the Sea Urchin

The study of fertilization in marine invertebrates, particularly in the sea urchin, an echinoderm, has been a favorite subject for scientists for many years [(21-23), reviewed by Epel (24)]. A great deal of our current knowledge of fertilization can be attributed to sea urchin studies. The reasons for this are the following. The sexes of sea urchins are separate and the gametes are easy to manipulate because fertilization is external. Sea urchins have a breeding season that lasts three to eight months, and the gametes of certain sea urchin species can be obtained in the laboratory all year round (25,26). A female can release up to 400 million eggs and a male as many as 100 billion sperm. The eggs can be fertilized in a glass beaker in sea water with near 100% success.

Egg

The mature sea urchin eggs are very large compared to the somatic cells. They are around 70 to 80 μm in diameter. The relatively enormous size of eggs is related to their role as depots for storage of various metabolites, macromolecules and organelles that accumulate during oogenesis. There are two acellular layers, vitelline layer and jelly coat, completely surrounding the egg plasma membrane. The vitelline layer, an extracellular glycoprotein coat 100-300 Å in thickness, is closely apposed to the plasma membrane (27-30). The numerous plasma microvilli are covered by the vitelline layer (6-8). The vitelline layer bears receptors for sperm on its external surface (34-37). The jelly coat is a transparent, amorphous material composed of glycoprotein and sulfated polysaccharides. The acrosome reaction-inducing activity resides in the jelly coat (38-41).

Sperm

The sea urchin sperm is small compared to the egg. It is only a few microns in diameter. It is a specialized cell which is adapted to reach, recognize, bind and fuse to eggs. The sea urchin sperm has a head, a midpiece, and a tail. The head contains the sperm nucleus and the acrosome granule which is a membrane bound vesicle subjacent to the plasma membrane. Immediately posterior to the acrosome granule is where the profilamentous actin is located. The latter is polymerized and forms the acrosome process upon induction by the egg jelly coat (42-44). The midpiece of the sperm contains a single mitochondrion. The tail is the flagellum at the rear.

The interaction of the egg jelly coat and the sperm plasma membrane is the very first event in sea urchin fertilization. This induces the sperm acrosome reaction and a number of changes involving ion movements and membrane permeability alterations (45-50). The acrosome reaction consists of the fusion of the sperm plasma membrane with the acrosome granule membrane (51), the exocytosis of the acrosome granule and the extrusion of the acrosome process (52). The sperm then attaches to the egg vitelline layer by the acrosome process. The next step is the fusion of the sperm and egg plasma membranes.

It has been shown that fertilization in sea urchins is in most cases species specific; insemination between species generally does not result in fertilization (40,41,53-56). Studies of the mechanism of fertilization have shown that both the induction of the acrosome reaction and the subsequent sperm adhesion contribute to the overall specificity of fertilization.

Acrosomal Reaction

The acrosome reaction is a necessary prerequisite to successful fertilization. It requires extracellular Ca^{++} and is accompanied by Ca^{++} uptake and

acid release (45-50). The release of acid from sperm is coincident with the monovalent cation Na^+ uptake, which suggests that Na^+ and H^+ movements are linked (47,49,57-58).

In nature, the acrosome reaction is induced by the jelly coat. The jelly coat can induce all the above physiological changes in the sperm. SeGall and Lennarz (40) isolated egg jelly coat from four species of sea urchins, *Arbacia punctulata*, *Strongylocentrotus purpuratus*, *Strongylocentrotus drobachiensis* and *Lytechinus variegatus*, and studied the specificity of the acrosome reaction in these species. The results showed that the specificity varies among these species. The acrosome reactions of *A. punctulata* and *S. drobachiensis* are the most specific in that they are induced only by their homologous jelly coat, while the *S. purpuratus* sperm reacts equally well with the homologous or *L. variegatus* jelly coat, but not with the *A. punctulata* jelly coat. The acrosome reaction of *L. variegatus* can be induced by the jelly coat of all four species. The purified fucose sulfate polysaccharides show the same specificity to induce the acrosome reaction as the whole jelly coat. Although they found that the chemical composition in the four jelly coats are very similar with approximately 20% sialoprotein and 80% fucose sulfate polysaccharide, the positions of glycosidic linkage and sulfation were different (41). The above data show in summary that the structural differences in the fucose polymers are responsible for the species-specific induction of the acrosome reaction.

Podell and Vacquier (59,60) recently purified a 210 kd glycoprotein from *S. purpuratus* sperm plasma membrane which binds to the living egg and isolated egg jelly in a species specific fashion. The discovery that antisera to this protein block both the Ca^{++} uptake and the Na^+/H^+ exchange normally induced in the sperm by the egg jelly coat supports the idea that it is the jelly coat receptor on the sperm plasma membrane. It will be interesting to isolate the jelly coat

component that binds to this receptor and study the interaction at the molecular level.

The Adhesion and the Binding of Eggs to Sperm

The acrosome reaction initiates the contact and the binding of the sperm to the egg which results in the fusion of the two plasma membranes. Summer and Hylander (56) found that the binding of the gametes plays an important role in species-specific fertilization. They studied the specificity of acrosome reactions using gametes from four echinoids and showed that while the morphologically normal acrosome reaction occurs in nine of eleven heterologous combinations, no sperm-egg binding or fertilization occurs. Moy and Vacquier have shown that bindin, a major acrosome protein in sea urchin sperm, is responsible for the adhesion and binding of the sperm to the egg (61). The evidence for this function of bindin is the following. First, the antibody raised against electrophoretically purified bindin shows that bindin coats the acrosome process and furthermore, it lies at the point of sperm-egg attachment (62). Sperm which have not undergone the acrosome reaction do not react with the bindin antisera. Secondly, the particulate bindin agglutinates the unfertilized eggs in a species specific fashion (63-65). Bindin isolated from *S. purpuratus* does not agglutinate the unfertilized eggs of *S. franciscanus* and vice versa.

The *S. purpuratus* bindin has been purified and partially sequenced. The purified bindin migrates as a 30,500 dalton molecular weight protein in a SDS-polyacrylamide gel. There is no detectable carbohydrate by the phenol-sulfuric acid or anthrone tests (62). Bindin from *S. franciscanus* has been purified and the limited N-terminal sequence compared with *S. purpuratus* showed 80% homology (66, and V. C. Vacquier, unpublished data). The significance of the heterogeneity cannot be determined at this stage.

Since bindin can agglutinate the unfertilized, dejellied eggs in a species-specific manner, there must exist a receptor in the egg vitelline layer that interacts with bindin. There is a good deal of evidence that supports this idea. Eggs treated with trypsin showed lower fertilization ability which could be attributed to the removal of either the bindin receptors on the vitelline layer or the glycoprotein in the jelly coat which induces the acrosome reaction (67,68). Furthermore, eggs treated with trypsin or the cortical granule protease are not agglutinated by bindin which implies that the receptors are cleaved by the protease (69). The result that sperm binds only to the external surface of isolated vitelline layers suggests that the receptors are localized on the outer surface of the vitelline layer (70). The isolation of the bindin receptor has been a difficult task due to its extraordinarily high molecular weight and its insolubility (71-73). Rossignol *et al.* isolated a sperm receptor on the egg surface of *S. purpuratus* which specifically binds to homologous sperm and inhibits fertilization (37). The carbohydrate-rich fragment of this glycoprotein was released by proteolytic digestion and it has been demonstrated that it can bind to sperm and inhibit fertilization. However, the binding of this carbohydrate-rich fragment of the receptor to the egg is not species specific, which implies that the intact glycoconjugate or the protein component is responsible for the species specific interaction of the receptor to the sperm.

Apparently the interactions of the jelly coat-jelly coat receptor and bindin-bindin receptor are not the only species specific events in fertilization. There are reports showing that some peptides associated with egg jelly of sea urchins are also species specific (74). Speract, a peptide with a sequence of Gly-Phe-Asp-Leu-Asn-Gly-Gly-Gly-Val-Gly, and resact, a peptide with a sequence of Cys-Val-Thr-Gly-Ala-Pro-Gly-Cys-Val-Gly-Gly-Gly-Arg-Leu-NH_2 were isolated from the egg jelly coat of *S. purpuratus* and *A. punctulata* respectively (75-78). Both

speract and resact increase the cGMP level, respiration and mobility of the homologous sperm at pH 7.8-8.0 (which is the pH of the normal sea water (79). The heterologous sperm is not affected by these peptides. Resact has been shown to act as a chemoattractant for *A. punctulata* sperm (80). The sperm receptor for speract has been identified and partially characterized (81). Whether the interaction of these peptides with their receptors play any role in species specific fertilization remains unknown.

While there are many studies focussing on the protein components involved in fertilization, few studies are directed at the genes that code for the components. This thesis reports the isolation and investigation of bindin cDNA and its pattern of gene expression in *S. purpuratus*. Briefly, there are several specific problems in fertilization which we propose to address using the bindin sequence. Bindin cDNA of *S. purpuratus* can be used to isolate bindin cDNA from other sea urchin species. Sequence comparisons will give insight into the mechanism of reproductive isolation at the molecular level. The structural comparison will be relevant to speciation and the evolution process, since bindin is certainly a critical molecule in speciation. The bindin gene is a tissue specific gene and the identification of the regulatory sequences and the other factors which mediate bindin expression will contribute to our knowledge of the mechanisms of developmental control in later sea urchin ontogeny. The bindin gene should also serve as a precise marker for examination of gonadal differentiation in the sea urchin, a little known process that has so far been described only in morphological terms (82).

References

1. L. H. Colwin and A. L. Colwin. J. Biophys. Biochem. Cytol. 10, 231 (1961).
2. A. L. Colwin and L. H. Colwin. J. Biophys. Biochem, Cytol. 10, 255 (1961).
3. A. L. Colwin and L. H. Colwin. J. Cell Biol. 19, 477 (1963).
4. L. H. Colwin and A. L. Colwin. J. Cell Biol. 19, 501 (1963).
5. M. H. Bernstein. Expl. Cell Res. 27, 197 (1962).
6. L. Franklin. J. Cell Biol. 25, 81 (1965)
7. J. J. Pasteels. Arch. Biol. (Liège) 76, 463 (1965).
8. B. A. Afzelius. Z. Zellforsch. Mikroskop. Anat. 42, 134 (1955).
9. J. C. Dan. Intern. Rev. Cytol. 5, 365 (1956).
10. P. C. Galtsoff and D. E. Philpott. J. Ultrastruct. Res. 3, 241 (1960).
11. J. J. Pasteels and E. de Harven. Arch. Biol. (Liège) 73, 445 (1962).
12. A. Tyler. As reported in studies on fertilization and early development. Eng. Sci. Mag. 27, 17, California Institute of Technology, Pasadena, California (1964).
13. M. H. Burgos and D. W. Fawcett. J. Biophys. Biochem. Cytol. 2, 223 (1956).
14. A. Anberg. Acta Obst. Gynecol. Scand. 36, Suppl. 2, 1 (1957).
15. D. W. Fawcett. Intern. Rev. Cytol. 7, 195 (1958).
16. E. Horstmann. Z. Zellforsch. Mikroskip. Ant. 54, 68 (1961).
17. R. Hadek. J. Ultrastruct. Res. 9, 110 (1963).
18. J. M. Bedford. J. Reprod. Fertility 7, 221 (1964).
19. R. G. Saacke and J. O. Almquist. Am. J. Anat. 115, 143 (1964).
20. L. Nicander and A. Bane. Z. Zellforsch. Mikroskip. Anat. 57, 390 (1962).
21. F. R. Lillie. Exp. Zool. 16, 524 (1914).
22. F. R. Lillie. Biol. Bull. (Woods Hole, Mass.) 28, 18 (1915).
23. J. Loeb. "The Organism as a Whole," University of Chicago Press (1916).
24. D. Epel. Curr. Top. Dev. Biol. 12, 186 (1978).

25. P. S. Leahy, B. R. Hough-Evans, R. J. Britten, and E. H. Davidson. J. Exp. Zool. 215, 7 (1981).

26. P. S. Leahy, T. C. Tutschulte, R. J. Britten, and E. H. Davidson. J. Exp. Zool. 204, 369 (1978).

27. J. Runnstrom. Adv. Morphol. 5, 221 (1966).

28. D. E. Chandler and J. Heuser. J. Cell. Biol. 84, 618 (1979).

29. P. Kidd. J. Ultrastruct. Res. 64, 402 (1978).

30. P. Kidd and D. Mazia. J. Ultrastruct. Res. 70, 58 (1979).

31. M. J. Tegner and D. Epel. J. Exp. Zool. 197, 31 (1976).

32. T. E. Schroeder. Dev. Biol. 64, 342 (1978).

33. D. Mazia, G. Schatten and R. Steinhardt. Proc. Natl. Acad. Sci. USA 72, 4469 (1975).

34. K. Aketa. Embryologia 9, 238 (1967).

35. K. Aketa, K. Onitake and H. Tsuzuki. Exp. Cell Res. 71, 27 (1972).

36. C. G. Glabe and W. J. Lennarz. J. Supramol. Struct. 15, 387 (1981).

37. D. P. Rossignol, B. J. Earles, G. L. Decker, and W. J. Lennarz. Dev. Biol. 104, 308-321 (1984).

38. K. Hotta, H. Hamazaki and M. Kurokawa. J. Biol. Chem. 245, 5434 (1970).

39. K. Hotta, M. Kurokawa and S. Isaka. J. Biol. Chem. 248, 629 (1973).

40. G. K. SeGall and W. J. Lennarz. Dev. Biol. 71, 33-48 (1979).

41. G. K. SeGall and W. J. Lennarz. Dev. Biol. 86, 87-93 (1981).

42. L. G. Tilney, S. Hatano, H. Ishikawa, and M. S. Mooseker. J. Cell Biol. 59, 109-126 (1973).

43. L. G. Tilney. J. Cell. Biol. 69, 51 (1976).

44. L. G. Tilney and M. S. Mooseker. J. Cell. Biol. 71, 402 (1976).

45. J. Dan. Biol. Bull. (Woods Hole, Mass.) 107, 335 (1954).

46. F. Collins and D. Epel. Exp. Cell. Res. 106, 211 (1977).

47. R. W. Schackmann and B. M. Shapiro. Dev. Biol. 18, 145 (1981).

48. G. L. Decker, D. B. Joseph and W. J. Lennarz. Dev. Biol. 53, 115 (1976).

49. R. W. Schackmann, E. M. Eddy and B. M. Shapiro. Dev. Biol. 81, 145 (1978).

50. J. C. Dan. Int Rev. Cytol. 5, 365 (1956).

51. L. H. Colwin and A. L. Colwin. In "Fertilization" Vol. I, 295, Academic Press, New York (1967).

52. J. Dan. Biol. Bull., 103, 54 (1952).

53. J. Loeb, R. King and A. R. Moore. Wilhelm Roux' Arch. Entwicklungsmach Organ. 150, 509 (1910).

54. A. R. Moore. Portugaliae Acta. Biol. Ser. A., 258 (1949).

55. J. Runnstrom. Symposia Soc. Exp. Biol. 6, 39 (1952).

56. R. B. Summer and B. L. Hylander. Exp. Cell. Res. 96, 63-68 (1975).

57. D. Nishioka and N. Cross. In "Cell Reproduction: Essays in Honor of Daniel Mazia," (E. R. Dirksen, D. M. Prescott and C. F. Fox, eds) 403, Academic Press, New York (1978).

58. B. M. Shapiro, R. W. Schackmann, C. A. Gabel, C. A. Foerder and M. J. Farrance. In "The Cell Surface: Mediator of Developmental Processes," (S. Subtelney and N. K. Wessells, eds.), 127, Academic Press, New York (1980).

59. S. B. Podell and V. D. Vacquier. J. Cell. Biol. 99, 1598 (1984).

60. S. B. Podell and V. D. Vacquier. J. Biol. Chem. Vol. 260, 1715-1718 (1985).

61. V. D. Vacquier and G. W. Moy. Proc. Natl. Acad. Sci. USA 74, 2456-2460 (1977).

62. G. W. Moy and V. D. Vacquier. Curr. Top. Dev. Biol. 13, 31 (1979).

63. C. G. Glabe and V. D. Vacquier. Nature 267, 836 (1977).

64. C. G. Glabe and W. J. Lennarz. J. Cell. Biol. 83, 595-604 (1979).

65. C. G. Glabe and W. J. Lennarz. J. Supramol. Struct. 15, 387 (1981).

66. V. D. Vacquier. The Cell Surface. Academic Press, 151-168 (1980).

67. K. Aketa, K. Onitake and H. Tsuzuki. Exp. Cell Res. 72, 27-32 (1972).

68. E. Schmell, B. J. Earles, C. Breaux, and W. J. Lennarz. J. Cell Biol. 72, 35-46 (1977).

69. V. D. Vacquier and G. W. Moy. In "Cell Reproduction: Essays in Honor of Daniel Mazia, "379-389, Academic Press (1978).

70. C. G. Glabe and V. D. Vacquier. J. Cell Biol. 75, 410-421 (1977).

71. E. Schmell and C. Breaux. Fed. Proc. 34, 642 (1975).

72. C. G. Glabe and V. D. Vacquier. Proc. Natl. Acad. Sci. USA 75, 881 (1978).

73. A. C. Lopo and V. D. Vacquier. Dev. Biol. 79, 325 (1980).

74. H. Ohtake. J. Exp. Zool. 198, 313 (1976).

75. J. R. Hansbrough and D. L. Garbers. J. Biol. Chem. 256, 1447 (1981).

76. N. Suzuki, K. Nomura, H. Ohtake and S. Isaka. Biochem. Biophys. Res. Commun. 99, 1238 (1981).

77. D. L. Garbers, H. D. Watkins, J. R. Hansbrough, A. Smith and K. S. Misono. J. Biol. Chem. 257, 2734 (1982).

78. N. Suzuki, H. Shimomura, E. W. Radany, C. S. Ramarao, G. E. Ward, J. K. Bentley and D. L. Garbers. J. Biol. Chem. 259, 14874 (1984).

79. N. Suzuki and D. L. Garbers. Biol. Reprod. 30, 1167 (1984).

80. G. E. Ward, C. J. Brokaw, D. L. Garbers and V. D. Vacquier. J. Cell. Biol., in press.

81. L. J. Dangott and D. L. Garbers. J. Biol. Chem. 259, 13712 (1984).

82. M. S. Houk, and R. T. Hinegardner. Biol. Bull. 159, 280-294 (1980).

CHAPTER 1

Sequence of mRNA Coding for Bindin, a Species Specific Sea Urchin Sperm Protein Required for Fertilization

Sequence of mRNA Coding for Bindin, a Species Specific Sea Urchin Sperm Protein Required for Fertilization

Boning Gao, Laurie E. Klein*, Roy J. Britten, and Eric H. Davidson

Division of Biology, California Institute of Technology, Pasadena, California 91125

*Present address: Laboratory of Molecular Biology, Madison, WI 53706

Summary. ***Bindin,*** a major protein of the sea urchin acrosome granule, mediates the species specific adhesion and binding of sperm to egg required to effect fertilization. We report the isolation and sequence of bindin cDNA clones prepared from *Strongylocentrotus purpuratus* testis RNA. The bindin gene appears to be productively expressed only in males and only in testes. Thus bindin mRNA is not detected in eggs, ovaries, early embryos or other adult tissues. The protein is produced from a 51 kilodalton (kd) precursor, which is subsequently processed to yield the mature 24 kd bindin protein.

Fertilization is a species specific process that in the sea urchin occurs by a series of steps. The initial physiological event as the sperm approaches the egg is the acrosome reaction, which is induced by the fucose sulfate of the egg jelly coat, and is mediated by sperm membrane receptors (*1,2*). This reaction results in exocytosis of the acrosomal granule, and the extrusion of the acrosome process. The sperm binds to the egg vitelline layer by means of the acrosome process, and the plasma membranes of the sperm and egg then fuse. The overall species specificity of fertilization depends on both the induction of the acrosome reaction and on the subsequent interaction that leads to the binding of the acrosome process to the egg surface (*3,4*). Thus in some interspecific combinations heterologous egg jelly does not induce the sperm acrosome reaction, while in others heterospecific acrosome reactions occur, but even when this is observed, fusion of heterospecific sperm and egg membranes fails to take place. Vacquier and associates (*5-8*) demonstrated that the major protein of the acrosome granule, which they named ***bindin,*** is the molecular species responsible for the recognition reaction by which the acrosome process is bound to a glycoprotein receptor embedded in the vitelline membrane of the egg. Partial purification and characterization of the sperm receptor glycoprotein has been reported (*8-10*). On activation of the sperm, bindin is exposed by the eversion of the acrosome

granule. Bindin molecules coat the external surface of the acrosomal process, and have been detected in the electron microscope by immunocytological methods at the exact site of the sperm-egg bond (*5*). Bindin has been purified and characterized chemically, as reviewed by Vacquier (*8*). It contains no detectable carbohydrate, and a partial amino acid sequence has been derived (*8*).

We report here the isolation and the sequence of bindin cDNA clones, the specificity of expression of the bindin gene, and the interesting result that bindin is made from a much larger precursor polypeptide.

Translated Sequence and Organization of Bindin mRNA

A λgt10 cDNA library (*11*) was constructed from testis poly(A) RNA, using random 8-12 nucleotide (nt) calf thymus DNA fragments as primers for the reverse transcriptase reaction. The *Strongylocentrotus purpuratus* males used for this preparation were obtained from an intertidal population in January, when spermatogenesis is occurring at maximal rate. About 40,000 recombinant clones were recovered. A synthetic 17-nt probe mixture including all possible sequences predicted by a known region of the bindin protein sequence (overlined (2) in Fig. 1b) was used to screen the testis library. A restriction map of one of the selected clones is shown in Fig. 1a, and the 1873 nt sequence of the bindin cDNA insert included in this clone is given in Fig. 1b. To determine how close to the 5' end of the mRNA the cDNA insert extends we carried out a primer extension experiment, utilizing a second synthetic oligonucleotide complementary to the region between positions 14 and 36 of the sequence shown (overline (1) in Fig. 1b). This oligonucleotide was reacted with testis poly(A) RNA, extended with reverse transcriptase, and the reaction products displayed electrophoretically on a sequencing gel (not shown). These experiments indicated that the 5' terminus of the mRNA is 50 nt beyond the 5' terminus of the cDNA clone.

The initial ATG codon occurring in the sequence shown in Fig. 1b (at position 59) is followed immediately by a stop codon TGA in the same reading frame. However, the second ATG signal to occur, at position 142, initiates a 1443 nt open reading frame. Thus, the bindin mRNA apparently has a 188 nt 5' leader sequence. Evidence to be presented elsewhere shows that in the bindin gene this leader sequence is interrupted by an intron at a position 17 nt 5' to the translational initiation site (at least one additional intron is also present within the coding region of the gene). Following the translation termination signal at position 1585 the cDNA insert continues for an additional 280 nt. This region of the 3' trailer sequence displays numerous termination codons, and it further extends for approximately 500 nt (including the poly(A) tract) which have not been sequenced, since the size of the mature mRNA is about 2500 nt.

The mature bindin protein contains 236 amino acids, according to the derived sequence shown. It begins at amino acid position 246 according to the match with the N-terminal bindin amino acid sequences obtained from the purified acrosome protein (*8*). Four blocks of amino acid sequence derived from mature bindin polypeptides 70, 80, 6 and 12 amino acids in length were available to us (8), and Fig. 1b shows that these could be aligned with the derived amino acid sequence of the cDNA clone with 97% fidelity. The bindin gene is present in only one copy per haploid genome, as shown in the set of genome blots reproduced in Fig. 2. Therefore the sequenced cDNA clone could not derive from a closely related gene, but must indeed represent the true bindin message. Since the N-terminal 235 amino acid polypeptide is not present in the mature protein these observations demonstrate that bindin is initially synthesized as a precursor polypeptide over twice the length of the mature acrosome protein. This polypeptide begins with a characteristic leader sequence (underlined in Fig. 1b), that includes many hydrophobic amino acids (*15*). There are four basic amino

acids at the junction between the 245 amino acid N-terminal polypeptide and the mature bindin proteins, which constitute typical cleavage sites for trypsin-like protein processing enzymes (*16*). These amino acids are marked by arrows in Fig. 1b (amino acids 242-245). Several additional basic amino acid pairs occur in the pro-bindin polypeptide. There is no evidence as to whether these also serve as protease cleavage sites. The amino acid compositions of the pro-polypeptide and the mature bindin protein differ in several respects. As shown in Fig. 3, the pro-polypeptide is characterized by a relatively high content of hydrophilic residues, 34%, compared to 14% for the mature bindin. Furthermore, all of the cysteine residues are located in the pro-polypeptide, which suggests the possibility of an internally cross-linked, compacted structure for this region. The mature bindin is on the other hand relatively rich in glycine and proline (particularly in the regions between amino acid 274 to 299, and 392 to the C terminus). High contents of glycine and proline are characteristics of rigid structural proteins. Two hydrophobic regions (indicated under the brackets in Fig. 3a), which are shorter than the typical transmembrane domain, might serve as sites for interaction with other hydrophobic sequences. The molecular weights calculated from the deduced protein sequences are 51 kd for the total prebindin molecule and 24 kd for the mature bindin protein.

The conclusion that bindin is derived by processing from a much larger precursor is supported by the immunoprecipitation experiment shown in Fig. 4. An anti-bindin antibody generously provided by V. Vacquier was reacted with the *in vitro* translation products of testis poly(A) mRNA, and displayed by SDS gel electrophoresis. As can be observed in Fig. 4 the polypeptide that reacts with the anti-bindin antibody has an apparent molecular weight over twice that of the mature bindin protein.

Developmental Expression of the Bindin Gene

As an initial enquiry into the specificity of expression of the bindin gene we carried out a search for bindin mRNA in various male and female sea urchin tissues. RNAs were extracted from eggs, ovaries, 40 hour (late gastrula) stage embryos, and from adult sea urchin coelomocytes, tube feet, lantern tissue, and intestines, and reacted by the RNA gel blot method with bindin cDNA probes. To control the condition of the extracted RNAs intact actin mRNAs were demonstrated in all of the preparations, as reported previously (*22*). Figure 5 shows that only the testis RNA contains detectable bindin message.

Discussion

The bindin gene is interesting from physiological, regulatory and evolutionary vantage points. An unexpected aspect of these observations that pertains to the physiological function of the proteins coded by the bindin message is the presence of the 245 amino acid N-terminal polypeptide. During spermatogenesis this polypeptide is evidently cleaved from the mature 236 amino acid bindin moiety. The acrosome granule is initially derived from the Golgi complex (*1*). Thus prebindin may be synthesized in the rough endoplasmic reticulum of the spermatocyte, transported to the Golgi complex, and later incorporated in the membrane-bound acrosome granule (*23*). Precedents in which mature proteins are initially synthesized as significantly larger precursors and are subsequently processed include certain cellular, viral, hormonal, and neuronal peptides (e.g., refs. 24-27). Among the functions proposed for such precursor regions are that they might facilitate transport of the protein to its ultimate intracellular destination (24); that they might be needed to ensure correct structural conformation (25-26); and that they are required for storage of the protein as an inactive form (*27*). An interesting analogy to the bindin case might be provided by the glycine- and proline-rich structural protein collagen. This protein is initially

synthesized as a soluble procollagen precursor, in which the cysteines of the propeptide form interchain disulfide bonds, the function of which is to stabilize the triple helical body of the protein (*26*). On the other hand, it is also possible that the N-terminal prebindin polypeptide has an independent function of its own. Though bindin is indeed the major protein of the isolated acrosome granule (*5*), other components may have been leached out during the isolation procedure. Several studies have indicated that on exocytosis the acrosome granule releases various enzymes (*28*) and at present it cannot be excluded that the N-terminal bindin polypeptide has this or some other functional acrosomal activity that is also required for fertilization.

We found the bindin gene to be productively expressed only in testis among the tissues studied, and hence it would appear that it is utilized only in males. It is clear from the qualitative experiments shown in Fig. *5* that bindin mRNA is not present in female gonads or in eggs. This is in contrast to the pattern of expression of the gene for the egg yolk protein vitellogenin. Though in all other animals so far investigated vitellogenin is synthesized exclusively in females, Shyu *et al.* (*29*) showed that in sea urchins vitellogenin is produced both in testes and in ovaries, as well as in intestine cells of both sexes. The observations reported here confirm that sexual determination in sea urchins does indeed involve differential expression of genes in male and female gonads, and they provide a specific molecular marker useful for further studies of this aspect of developmental gene regulation. Other evidence (*30*) demonstrates that there are also H1 and H2B histone genes that are expressed specifically in testis, the products of which are found only in sperm nuclei. Thus there is evidently a battery of male-specific genes, that are activated in the preparation of the sperm, and are otherwise quiescent.

Change in the primary sequence of essential coding elements of the bindin gene could have played a crucial role in the evolutionary processes leading to speciation. Thus preference in the bindin:bindin-receptor interaction within a population isolate would lead directly to isolation of its gene pool. In contemporary sea urchins this interaction indeed apparently functions as a barrier to gene flow between species such as *S. purpuratus* and *S. franciscanus*. Viable hybrid embryos can be formed between these species (*31*), and rare naturally occurring adult hybrids of these species have been reported (*32*). Thus the interspecific fertilization barrier rather than developmental incompatibility is probably a limiting factor in preserving the genetic separateness of these species. *S. purpuratus* sperm bindin fails to react with the receptor on *S. franciscanus* eggs and *vice versa* (*4,6,8*). The bindin proteins of these species are of the same molecular weight, and the sequence of the initial 73 amino acids of the mature *S. franciscanus* bindin has been obtained (*8,33*). For the first 40 residues of the mature protein the amino acid sequences are 80% homologous. However between residues 31 and 59 the *S. franciscanus* sequence consists of three imperfect tandem repeats of a 10 amino acid sequence element (*34*), only the first of which is present in the *S. purpuratus* bindin at this location. Elsewhere in the *S. purpuratus* bindin molecule, sequences homologous to the first 5 amino acids of this repeat occur in 8 additional locations. The repetition of this sequence motif might suggest that it is functionally important, and the differing organization of the same or similar repeats in the bindin molecules of the two species could contribute to the molecular basis for their different functional recognition specificities. The extent and nature of the distinctions in the primary sequences of the two proteins will be clarified when the *S. franciscanus* bindin gene is cloned and fully sequenced (in progress). The development of methods for inserting genes into the sea urchin genome (*35*) provides the means to obtain more exact

knowledge of the functional molecular changes that during evolution resulted in the genetic isolation of the populations ancestral to these two species.

References and Notes

1. J. C. Dan, in *Fertilization*, Vol. I, C. B. Mertz and A. Monroy, Eds. (Academic Press, New York, 1967), p. 237.
2. R. G. Summers, B. L. Hylander, L. H. Colwin, A. L. Colwin, *Am. Zool.* **15**, 523 (1975); S. B. Podell and V. D. Vacquier, *J. Biol. Chem.* **260**, 2715 (1985).
3. R. G. Summers and B. L. Hylander, *Exp. Cell. Res.* **96**, 63 (1975); G. K. Segall and W. J. Lennarz, *Dev. Biol.* **71**, 33 (1979).
4. C. Glabe, M. Buchalter, W. J. Lennarz, *Dev. Biol.* **84**, 397 (1981).
5. G. W. Moy and V. D. Vacquier, *Curr. Topics Dev. Biol.* **13**, 31 (1979).
6. C. G. Glabe and V. D. Vacquier, *Nature* **267**, 30 (1977).
7. V. D. Vacquier and G. W. Moy, *Proc. Natl. Acad. Sci. USA* **74**, 2456 (1977).
8. V. D. Vacquier, in *The Cell Surface, Mediator of Developmental Processes*, N. K. Wessells, Ed. (Academic Press, New York, 1980), p. 151. Additional unpublished bindin protein sequence data were made available to us by V. Vacquier.
9. D. P. Rossignol, B. J. Earles, G. L. Decker, W. J. Lennarz, *Dev. Biol* **104**, 308 (1984).
10. M. Acevado-Duncan and E. Carroll, *Gamete Research*, in press.
11. C. J. Watson and J. F. Jackson, in *DNA Cloning: A Practical Approach*, D. Glover, Ed. (IRL Press, Oxford, 1984), p. 79. We followed the protocol given in this reference for the synthesis of the double strand cDNA. Linker ligation, size selection of the cDNA, and the following procedures are as described by T. V. Huynh, R. A. Young, R. W. Davis in the same book, p.49. The average length of the insert obtained was 1.2 kb.
12. F. Sanger, S. Nicklen and A. R. Coulson, *Proc. Natl. Acad. Sci. USA* **74**, 5463 (1977).

13. J. Messing, *Meth. in Enzymol.* **101**, 20 (1983).

14. S. Henikoff, *Gene* **28**, 351-359 (1984).

15. M. E. E. Watson, *Nucleic Acids Res.* **12**, 5145 (1984).

16. M. Bergmann, and J. S. Fruton, *Adv. Enzymol.* **1**, 63 (1941).

17. J. W. Posakony, R. H. Scheller, D. M. Anderson, R. J. Britten, E. H. Davidson, *J. Mol. Biol.* **149**, 41 (1981).

18. R. J. Britten, A. Cetta, E. H. Davidson, *Cell* **15**, 1175 (1978); J. J. Lee, R. J. Shott, S. J. Rose III, T. L. Thomas, R. J. Britten, E. H. Davidson, *J. Mol. Biol.* **172**, 149 (1984).

19. S. W. Kessler, *J. Immunol.* **115**, 1617 (1975).

20. F. L. Greenwood, W. M. Hunter, J. S. Glover, *Biochemical J.* **89**, 184 (1963).

21. V. D. Vacquier, *Anal. Biochem.* **129**, 497 (1983).

22. R. J. Shott, J. J. Lee, R. J. Britten, E. H. Davidson, *Dev. Biol.* **101**, 295 (1984).

23. M. H. Burgos and D. W. Fawcett, *J. Biophysic Biochem. Cytol.* **1**, 287 (1955); M. H. Burgos and D. W. Fawcett, *J. Biophysic Biochem. Cytol.* **2**, 223 (1956).

24. A. Viebrock, A. Perz, W. Sebald, *EMBO J.* **1**, 565 (1982).

25. R. E. Chance, R. M. Ellis, W. W. Bromer, *Science* **161**, 165 (1968).

26. P. Bornstein, *Ann. Rev. Biochem.* **43**, 567 (1974).

27. B. Kassell and J. Kay, *Science* **180**, 1022 (1973).

28. A. E. Levine, K. A. Walsh, E. J. B. Fodor, *Dev. Biol.* **63**, 299 (1978); J. G. Csernansky, M. Zimmerman, W. Troll, *Dev. Biol.*, **70**, 283 (1979); A. F. Conway and C. B. Metz, *J. Exp. Zool.* **198**, 39 (1976).

29. A. Shyu, R. A. Raff, T. Blumenthal, *Proc. Natl. Acad. Sci USA*, in press.

30. M. Strickland, W. N. Strickland, W. F. Brandt, and C. Von Holt, *Eur. J. Biochem.* **77**, 263 (1977); W. N. Strickland, M. Strickland, P. C. de Groot, C. Von Holt, B. Wittmann-Liebold, *Eur. J. Biochem.* **104**, 559 (1980); M. C. Busslinger, A. Barberis. *Proc. Natl. Acad. Sci. USA* **82**, 5676 (1985).

31. J. Loeb, R. King, A. R. Moore, *Wilhelm Roux' Arch Entwicklungsmech. Organ.* **150,** 509 (1910); R. R. Chaffee and D. Mazia, *Dev. Biol.* **7,** 502 (1963); D. Barrett and G. M. Angelo, *Exp. Cell Res.* **57,** 159 (1969).

32. E. Swan, *Evolution* **7,** 269 (1953).

33. Unpublished data from V. D. Vacquier.

34. From unpublished data of V. Vacquier, and Fig. 1b of this paper, the two sequences are as follows:

```
Sp 1   TyrValAsnThrMetGlyTyrProGlnAlaMetSerProGlnMetGlyGlyValAsnTyrGlyGlnProAlaGln
Sf      - Gln - GlnGlyAsn -  -  -  -  - Asn -  - Ser -  -  -  -  -  -  -  -  -  -

Sp 26  GlnGlyTyrGlyAlaGlnGlyMetGlyGlyProValGlyGlyGlyProMetGlyGlyProProGlnPheGlyAla
Sf      -  -  -  -  - [-  -  -  -  -  AlaPhe -  - -][GlnGlyMet - GlyAlaVal *  - Gly]

Sp 51  LeuProProGlyGlnAlaAspThrAspPheGlySerSerSerSerSerValAspGlyGlyAspThrThr
Sf     [GlnGlyMet - Gly - ValGlyGlyGly] * Phe * AlaPhe - ProGlyGluAlaGluAlaAsp
```

The bracket indicates the ten amino acid motif that is repeated in S.f. sequence. Amino acids that were not determined are marked with an asterisk.

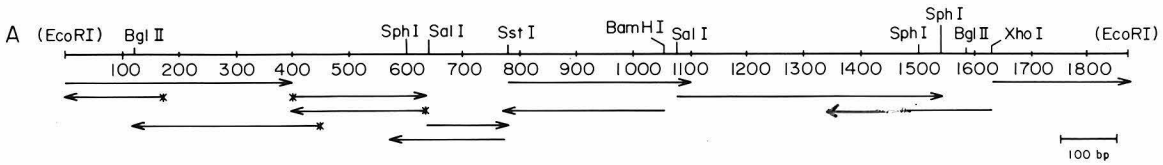
35. Reviewed in E. H. Davidson, C. N. Flytzanis, J. J. Lee, J. J. Robinson, S. J. Rose, III, H. M. Sucov, *Cold Spring Harbor Symp. Quant. Biol.* **50,** 321 (1985).

36. J. Kyte and R. F. Dootlittle, *J. Mol. Biol.* **157,** 105 (1982).

37. We thank Dr. V. Vacquier for the generous gift of bindin protein and bindin antibody, and for allowing us to cite unpublished data. We also thank Dr. S. J. Horvath for providing synthetic probes and primer, and Mr. J. A. Ferrier for technical assistance. B. Gao was supported by California Foundation for Biochemical Research. Research supported by NSF grant DCB8502191.

Fig. 1 Restriction map and structure of bindin cDNA. (a) Cleavage sites and sequencing strategy for bindin cDNA. The lengths of the arrows indicate the direction and number of nucleotides for which sequence was determined by the dideoxynucleotide chain termination method (*12*, *13*). The fragments marked with asterisks were generated by controlled deletion with Exonuclease III (*14*). The *Eco*RI sites at the ends of the cDNA derive from the *Eco*RI linkers used to construct the cDNA library. This clone was selected from the testis cDNA library by screening with a 17 nt long probe mixture (see text and b). The probe DNA was labeled by the kinase reaction to a specific activity of 3×10^8 cpm/μg, and hybridized with filters bearing plaque lifts in an aqueous medium at 42°C. (b) Nucleotide sequence of bindin cDNA and predicted amino acid sequence of the bindin precursor. Nucleotides are numbered from the 5' terminus. The amino acid sequence derived from the nucleotide sequence is numbered from the first methionine of the prebindin polypeptide. Some regions of the mature bindin protein sequence have been determined (*8*). Dashes indicate residues that are identical with the predicted amino acid sequence, while the italicized residues differ from the predicted protein sequence. The underlined sequence indicates a hydrophobic sequence of amino acids that display the characteristics of a leader sequence. Arrows indicate basic amino acids that are potential cleavage sites for trypsin-like enzymes. The oligonucleotides utilized for the primer extension experiment and for the isolation of this clone (see text) were derived from the regions of sequence indicated by overlines (1) and (2) respectively.

A



B

```
1                    21                   41                   61                   81                  101
|                    |                    |                    |                    |                    |
                     [1]
TAG CCA GAA AAT ACG CTG AAG TGA AAA GGA CAG ACG AAG CAT TCT GCT ACA TTT CTT CAT GTG AAA AAC TCG ATT TCG AAG ATT AGC AGT TCT TTT TTT CAC TCG TGC ACA TTT TTA TTA

121                  141                  161                  181                  201                  221
|                    |                    |                    |                    |                    |
CAA GAT CTA CAT TTA AGC ATC ATG GGT TTC CAT CAA ATT TTA GTC ACT GTA GTT GCC CTC GCT TTA GCC TCT GTC AGA GCC GAG TTC CCA TCC CGT ACC GAC TCC CCT ACT GAC TGC CCC
                         Met Gly Phe His Gln Ile Leu Val Thr Val Val Ala Leu Ala Ser Val Arg Ala Glu Phe Pro Ser Arg Thr Asp Ser Pro Thr Asp Cys Pro
                                             10                   20                                        30

241                  261                  281                  301                  321                  341
|                    |                    |                    |                    |                    |
GAA GCA GAT CAA GGG TGC TGG TGT CGC GGC TCC TTC GCT CAG TGC TGG AGA ACG TAT GAA GAG GCA GGA ATG ACA GGA GAA ATT GGA AAC AGA ATT ACA AAG CTA GAC TTA TTG TAT CAG
Glu Ala Asp Gln Gly Cys Trp Cys Arg Gly Ser Phe Ala Gln Cys Trp Arg Thr Tyr Glu Glu Ala Gly Met Thr Gly Glu Ile Gly Asn Arg Ile Thr Lys Leu Asp Leu Leu Tyr Gln
                    40                   50                   60                   70

361                  381                  401                  421                  441                  461
|                    |                    |                    |                    |                    |
CCA AGT GAA GAG ATT GTG ACC TAC ATA AGA CGT AGT AGC GCC ATG AGG GAA TTA AGA ATT TCA GAG GAT GGC GTT AGC TTA GAT TGT TCT TGC GAC CTC ATC TAT GCC CTC GAT GAC AAA
Pro Ser Glu Glu Ile Val Thr Tyr Ile Arg Arg Ser Ser Ala Met Arg Glu Leu Arg Ile Ser Glu Asp Gly Val Ser Leu Asp Cys Ser Cys Asp Leu Ile Tyr Ala Leu Asp Asp Lys
                    80                   90                   100                  110

481                  501                  521                  541                  561                  581
|                    |                    |                    |                    |                    |
CAT GTA ACA CTC GTT GAC CAG GCC GAG CTG ACG TTC AGT AAC TGC CAG CAG AGG GGA TGG CCT CGT GAT AGC ATG ACT GCC AGA TCC TTC GTT AAT CGC TGC CAT GTT AGT CGC ATG CAA
His Val Thr Leu Val Asp Gln Ala Glu Leu Thr Phe Ser Asn Cys Gln Gln Arg Gly Trp Pro Arg Asp Ser Met Thr Ala Arg Ser Phe Val Asn Arg Cys His Val Ser Arg Met Gln
                    120                  130                  140                  150

601                  621                  641                  661                  681                  701
|                    |                    |                    |                    |                    |
GAT GGC GAT TTA CGA AAG CGG AGA GAA TCT GAG GAT GTC GAC GAC GAT GAT GTC AGC AAA AGG GCA AGT CCT CGT AAG GGG GAC GAA CCA GCT GGT CAT ACT CTT AAG GAC CTC GCC CCC
Asp Gly Asp Leu Arg Lys Arg Arg Glu Ser Glu Asp Val Asp Asp Asp Asp Val Ser Lys Arg Ala Ser Pro Arg Lys Gly Asp Glu Pro Ala Gly His Thr Leu Lys Asp Leu Ala Pro
                    160                  170                  180                  190

721                  741                  761                  781                  801                  821
|                    |                    |                    |                    |                    |
CAA AAC ACC AAT CAT CTT GTT AGC ATC GAT GGT GCC GAC AAA CAT CCA GCC GAC GAG CTC GTC AAC TTC ATC TCG GGG CAT TCT CCG ACT CGA CGT GCG ACA GAC AAC GAT GCC GCA GTT
Gln Asn Thr Asn His Leu Val Ser Ile Asp Gly Ala Asp Lys His Pro Ala Asp Glu Leu Val Asn Phe Ile Ser Gly His Ser Pro Thr Arg Arg Ala Thr Asp Asn Asp Ala Ala Val
                    200                  210                  220                  230

841                  861                  881                  901                  921                  941
|                    |                    |                    |                    |                    |
TCT GAC GAT TCG AAA AGA GGA GCG CGT AAG AAG CGT TAC GTC AAC ACG ATG GGC TAC CCC CAA GCC ATG AGT CCG CAA ATG GGA GGT GTT AAC TAC GGA CAA CCA GCG CAA CAA GGG TAC
Ser Asp Asp Ser Lys Arg Gly Ala Arg Lys Lys Arg Tyr Val Asn Thr Met Gly Tyr Pro Gln Ala Met Ser Pro Gln Met Gly Gly Val Asn Tyr Gly Gln Pro Ala Gln Gln Gly Tyr
                    240  +   +   +   +        250                  260                  270

961                  981                  1001                 1021                 1041                 1061
|                    |                    |                    |                    |                    |
GGC GCT CAA GGA ATG GGT GGA CCA GTC GGA GGT GGA CCG ATG GGT GGA CCT CCA CAA TTT GGT GCG TTA CCT CCA GGG CAG GCA GAT ACG GAT TTT GGA TCC AGC TCC TCC AGC GTC GAC
Gly Ala Gln Gly Met Gly Gly Pro Val Gly Gly Gly Pro Met Gly Gly Pro Pro Gln Phe Gly Ala Leu Pro Pro Gly Gln Ala Asp Thr Asp Phe Gly Ser Ser Ser Ser Val Asp
                                                                                  Ser         Glu                                                  Val
                    280                  290                  300                  310

1081                 1101                 1121                 1141                 1161                 1181
|                    |                    |                    |                    |                    |
GGA GGT GAT ACA ACG ATC AGC GCC AGA GTT ATG GAC GAC ATC AAG GCG GTG CTC GGT GCC ACG AAG ATT GAC CTA CCA GTA GAC ATC AAC GAC CCT TAC GAT CTT GGA CTC CTG CTA CGC
Gly Gly Asp Thr Thr Ile Ser Ala Arg Val Met Asp Asp Ile Lys Ala Val Leu Gly Ala Thr Lys Ile Asp Leu Pro Val Asp Ile Asn Asp Pro Tyr Asp Leu Gly Leu Leu Leu Arg
Glu Lys                                                                                                                                                                  
                    320                  330                  340                  350

1201                 1221                 1241                 1261                 1281                 1301
|                    |                    |                    |                    |                    |
CAC CTC AGG CAT CAC TCC AAC CTT CTA GCG AAC ATC GGT GAC CCT GCG GTT CGG GAA CAG GTC CTA AGC GCG ATG CAG GAG GAC GAA GAA GAG GAG GAA GAA GAT GCA GCA ACC GGC GCC
His Leu Arg His His Ser Asn Leu Leu Ala Asn Ile Gly Asp Pro Ala Val Arg Glu Gln Val Leu Ser Ala Met Gln Glu Glu Glu Glu Glu Glu Glu Glu Asp Ala Ala Thr Gly Ala
                    360                  370                  380                  390

1321                 1341                 1361                 1381                 1401                 1421
|                    |                    |                    |                    |                    |
CAA CAA GGC GTC TTG AAC GGG AAC GCA CCA GGC CAA GCA GGC TTC GGT GGC GGA GGT GGA GGA GGG GCC ATG ATG AGT CCT CAA CAG ATG GGC GGA CAA CCT CAA GGG ATG ATC GGA CAA
Gln Gln Gly Val Leu Asn Gly Asn Ala Pro Gly Gln Ala Gly Phe Gly Gly Gly Gly Gly Gly Gly Ala Met Met Ser Pro Gln Gln Met Gly Gly Gln Pro Gln Gly Met Ile Gly Gln
                    400                  410                  420                  430

1441                 1461                 1481                 1501                 1521                 1541
|                    |                    |                    |                    |                    |
CCT CAA GGG ATG GGC TTT CCA CAT GAA GGG ATG GGC GGA CCA CCC CAA GGG ATG GGC ATG CCA CAT CAA GGG ATG GGC GGA CCA CCT CAA GGG ATG GGC ATG CCA CCT CAA GGG CAA CCG
Pro Gln Gly Met Gly Phe Pro His Glu Gly Met Gly Gly Pro Pro Gln Gly Met Gly Met Pro His Gln Gly Met Gly Gly Pro Pro Gln Gly Met Gly Met Pro Pro Gln Gly Gln Pro
                    440                  450                  460                  470

1561                 1581                 1601                 1621                 1641                 1661
|                    |                    |                    |                    |                    |
TAC GGT CAG GGT TAT CTA CAG GGG TAG ATC TAA CTA GAT ATA GAC GCT AAT TTG GTG TAT ATC GGG GTG ATC TCG AGA GCG TTT GCA AAA TAT ATT CTG AGA GTG GAG TTC TTC CTT TAG
Tyr Gly Gln Gly Tyr Leu Gln Gly ---
                    480

1681                 1701                 1721                 1741                 1761                 1781
|                    |                    |                    |                    |                    |
AAC ACG AGT ATT TGA ATT GAC TTT TCA ATT TTC TTG TTC GGT GAG GTG CAA CAT CAT TAT AGC TAT TTA GTT AAA AAA AAA CGT TTC TTT TTG GCA AAT TTG ATA CAC GTA TTC ATT GTA

1801                 1821                 1841                 1861
|                    |                    |                    |
AAA AAA CCA ACA ACA CCG ACA TGA AAT ATA AGC GAC GAT AAA GAG AGA GAA AAA GAG GGG AGG CCC AGA CTC G
```

Fig. 2. The bindin gene is a single copy sequence. Genome blots were prepared with sperm DNAs from four different male *S. purpuratus* individuals digested with *Hin*dIII, and reacted with bindin ^{32}P-cDNA probes. Hybridization was at 68°C in a medium containing 0.75 M Na^+, 5X Denhart's solution, 50 μg/ml denatured, sheared calf thymus DNA, and 2 mM phosphate buffer, pH 6.8, and equivalent criterion conditions were applied for washing. The molecular weight markers were λ DNA digested with *Hin*dIII. As is characteristic for sea urchin genes (*17*,*18*) the flanking sequences are polymorphic, and the genome blot reveals a number of different allelic variants. These are indicated (A-E) to the right of each lane. Thus lane 1 displays alleles B+D; lane 2 appears to be homozygous for allele C; lane 3 displays alleles A+E; lane 4, alleles C and D. This pattern indicates the bindin gene to be single copy since each DNA preparation displays a unique combination of two of the five alleles observed.

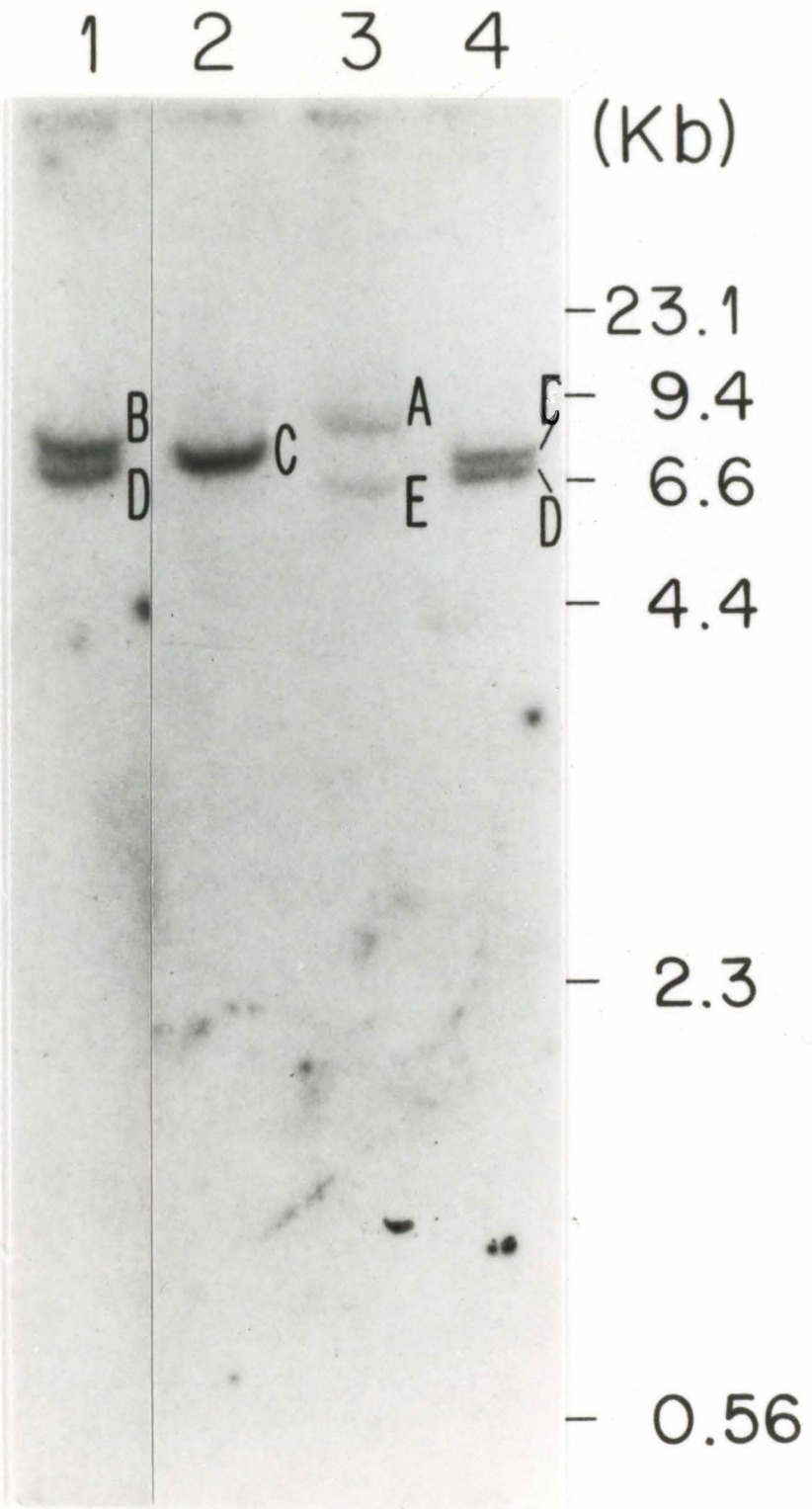

1
2
3
4
(Kb)
23.1
9.4
6.6
4.4
2.3
0.56
B
D
C
A
E
D

Fig. 3. (a) Hydrophobicity plot for the bindin protein sequence, obtained by the program of Kyte and Doolittle (36), using a search length of nine amino acids. Positive values indicate hydrophobicity and negative values hydrophilicity. The dotted vertical line denotes the putative leader sequence cleavage site. The N terminal of the mature bindin is indicated by the solid vertical line. The brackets demarcate two regions which are relatively hydrophobic. (b) Distribution of selected amino acids in the bindin protein. Each vertical line indicates one amino acid. (1) Acidic amino acids including aspartic acid and glutamic acid. (2) Basic amino acids, including histidine, arginine and lysine. (3) Glycine. (4) Proline. (5) Cysteine. The amino acids are numbered from the N terminal of the pro-bindin.

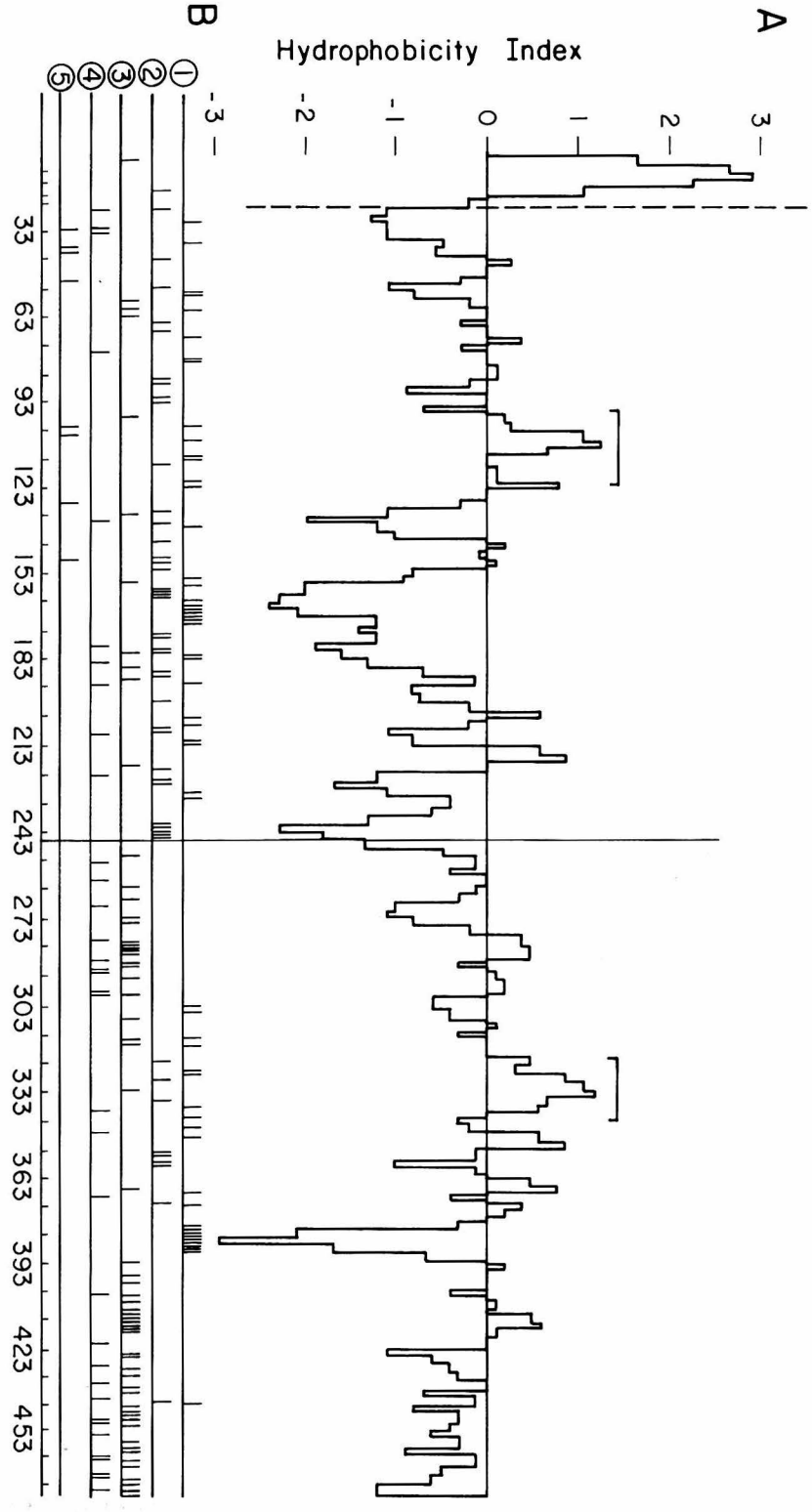

A
Hydrophobicity Index
3
2
1
0
-1
-2
B
-3
1
2
3
4
5
33
63
93
123
153
183
213
243
273
303
333
363
393
423
453

Fig. 4. Immunoprecipitation of *in vitro* translation products from testis poly(A) RNA with anti-bindin antibody. RNA was isolated from testes collected early in the spawning season. Following dissection, the testes were pelleted by centrifugation in a table-top centrifuge for 5 minutes. The procedure utilized for extracting RNA has been described in detail elsewhere (*17*). The testis poly(A) RNA was translated *in vitro* in a commercial rabbit reticulocyte lysate in the presence of 1 μCi/μl of ^{35}S-methionine (1200 Ci/mmol), at 30°C for 1 hour. The lysate was then diluted 1:10 with immunoprecipitation buffer containing 0.16 M NaCl, 20 mM Tris, pH 7.4, 0.1% deoxycholate and 1% Triton X-100. Rabbit anti-bindin serum or normal rabbit serum was added and the mixture was incubated at 4°C for 4 hours. The antibody complex was precipitated with fixed *Staphylococcus aureus* (Cowan I strain) (*19*) and fractionated over a 12.5% SDS-polyacrylamide gel. Lane A: *In vitro* translation with testis RNA, immunoprecipitated with normal serum; Lane B: *In vitro* translation with testis RNA, immunoprecipitated without antibody; Lane C: *In vitro* translation with testis poly(A) RNA, immunoprecipitated with bindin antibody, but in the presence of 5×10^{-3} μg of unlabeled mature bindin protein; Lane D: *In vitro* translation with testis poly(A) RNA, immunoprecipitated with bindin antibody in the presence of 5×10^{-3} μg of unlabeled ovalbumin; Lane E: Testes poly(A) RNA *in vitro* translation product; Lane F: Egg poly(A) RNA *in vitro* translation product; Lane G: *In vitro* translation with egg poly(A) RNA, immunoprecipitated with normal serum; Lane H: *In vitro* translation with egg poly(A) RNA, immunoprecipitated with bindin antibody. The apparent molecular weight of the polypeptide precipitated in the sample displayed in lane D is ~60 kd (indicated by an arrow at left), rather than the 51 kd calculated from the sequence shown in Fig. 1b. This is probably due to anomalous migration of this protein in SDS-polyacrylamide gels, relative to standards. Thus the mature bindin protein was reported to migrate in gels as a

30.5 kd protein (*8*), while HPLC measurements yielded a mass of 25 kd (*8*), close to the value deduced from the sequence of the mature protein, which demonstrates that the antibody utilized in (D) immunoprecipitates authentic bindin proteins. Lane I: Partially purified iodinated (*20*) preparation of acrosomal protein (*21*) the major component of which is bindin. After iodination this protein migrates more slowly than does native mature bindin (see arrow at left margin of (A)). Lane J: Immunoprecipitated component of preparation shown in I.

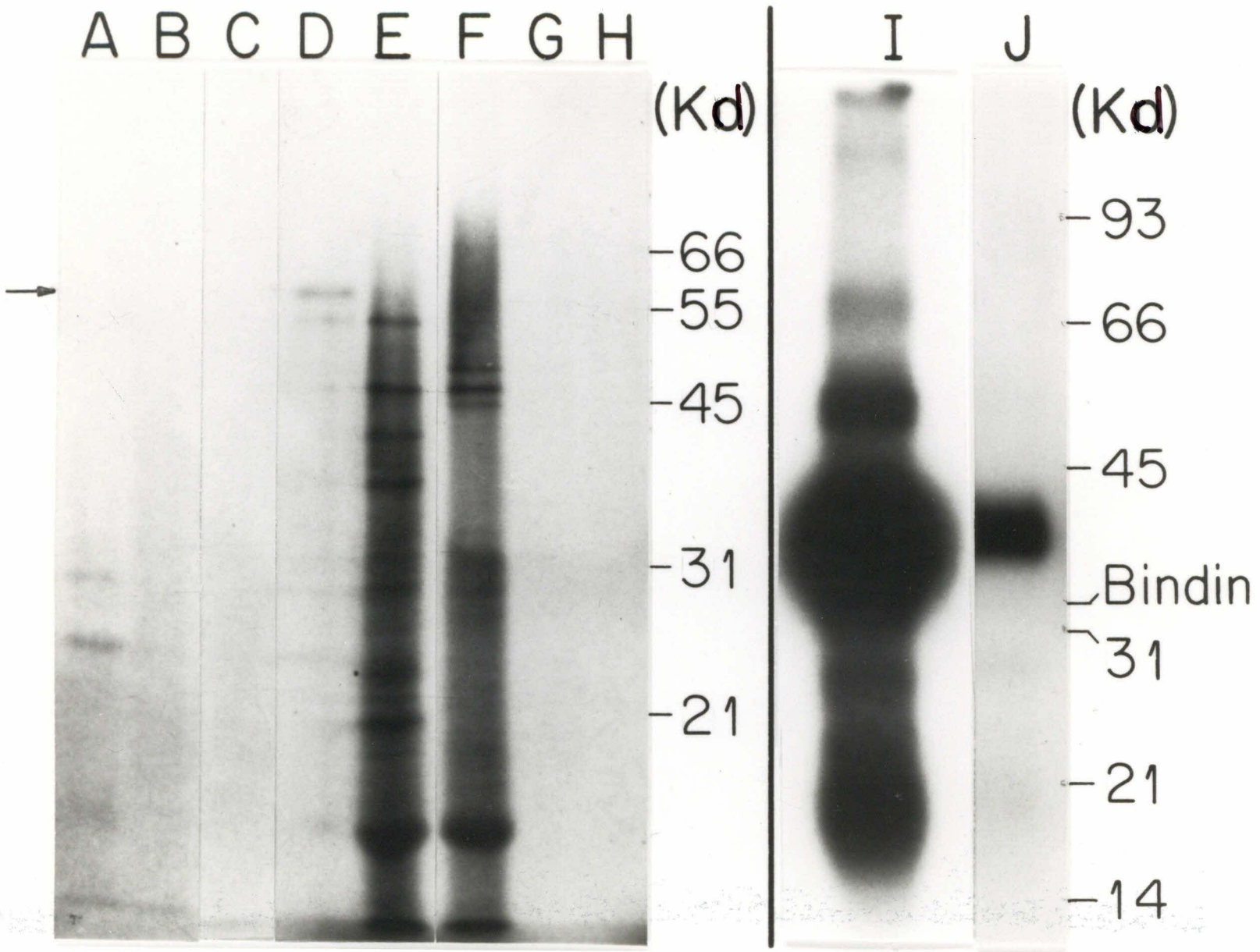
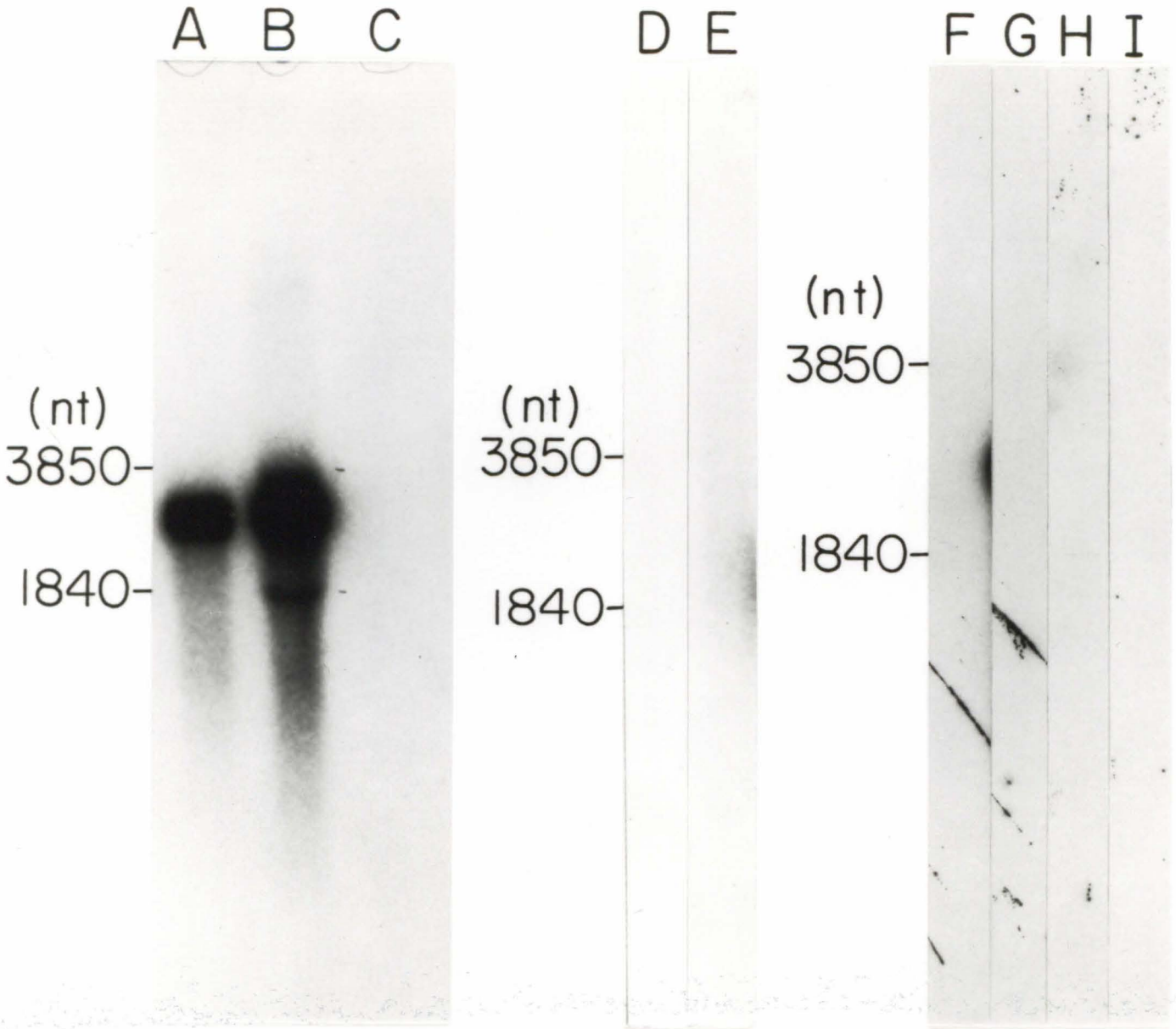

A B C D E F G H
(Kd)
66
55
45
31
21
I J
(Kd)
93
66
45
Bindin
31
21
14

Fig. 5. Bindin mRNAs in various sea urchin tissues and cell types. RNAs were displayed on a 1% formamide denaturing agarose gel and transferred to a nitrocellulose filter. The probe was the 1.4 kb *Bg*lII fragment of bindin cDNA (Fig. 1), labeled by nick translation to a specific activity of 3×10^8 cpm/μg. Hybridization was carried out in 50% formamide, 5X SSC, 1X Denhardt's solution, 2 mM phosphate buffer containing 50 μg/ml denatured and sheared calf thymus DNA at 42°C, for 16 hours. The filter was washed with 5X SSC, 2X SSC and 1X SSC successively, at 42°C. The autoradiograph was exposed with an intensifying screen. Lane A: 1 μg of testis poly(A) RNA exposed for 3 hrs. Lanes B-I, exposed for 5 hrs; Lane B: 15 μg testis total RNA; Lane C; 15 μg total RNA from 40 hour embryos; Lane D: 3 μg ovary poly(A) RNA; Lane E: 2 μg egg poly(A) RNA; Lane F: 1 μg coelomocyte poly(A) RNA; Lane G: 1 μg intestine poly(A) RNA; Lane H: 15 μg tube feet total RNA; Lane I: 1 μg lantern poly(A) RNA.

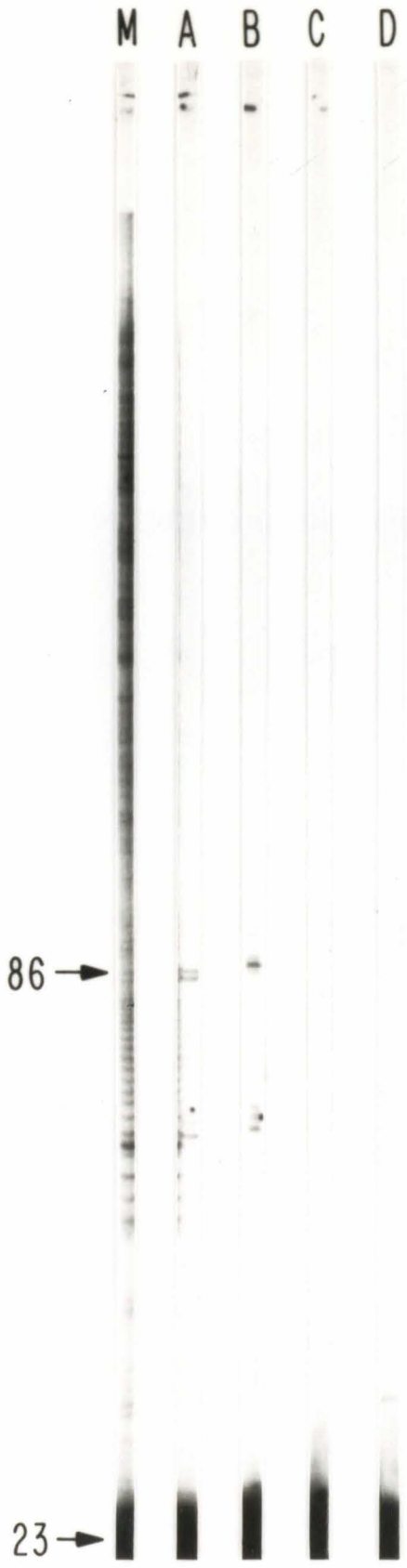

A B C
D E
F G H I
(nt)
3850-
1840-
(nt)
3850-
1840-
(nt)
3850-
1840-

APPENDIX I to CHAPTER 1

Appendix I

Primer Extension Experiment to Determine the 5' End of Bindin mRNA.

The single strand primer 5'CGTCTGTCCTTTTCACTTCAGCG corresponds to bindin cDNA from base pair 14 to 36. The primer extension experiment was performed in the presence of A. 1 μg testes poly(A)$^+$ RNA; B. 16 μg testes total RNA; C. 1 μg ovary poly(A)$^+$ RNA; D. No added RNA. Lane M is the molecular weight marker derived from a sequencing reaction. The number at left indicates the length of the band in nucleotides.


M
A
B
C
D
86
23

APPENDIX II to CHAPTER 1

Appendix II

Experimental Procedures

The Isolation of Sea Urchin Testis RNA.

Five male sea urchins were injected with 0.5 M KCl to release the mature sperm. The testes were removed and cut into small pieces. The tissue was rinsed with MPFSW (Millipore filtered sea water) several times to get rid of the mature sperm. The tissue was spun in a clinical centrifuge for 3 minutes at maximal speed. 10 ml of testis tissue was used for each RNA preparation. 33.6 g of Urea, 40 ml 2X lysis buffer (2X buffer is 30 mM NaOAC pH 5.5, 20 mM EDTA, 30 mM EGTA, 2% SDS, and 20 μg/ml polyvinyl sulfate) 40 mg protease K and dH_2O were added to a final volume of 80 ml. This solution was shaken vigorously in a 250 ml plastic bottle until all the tissues were dissolved. The solution was incubated at room temperature for 30 minutes. The following ultracentrifugation of RNA through a 5.7 M CsCl cushion has been described elsewhere (1). Poly(A)$^+$ RNA-enriched testis RNA was prepared by oligo (dT) cellulose chromatography (2).

The Construction of the λgt10 Random Primed Testis cDNA Library.

First strand cDNA was synthesized from testis poly(A)$^+$ RNA by using calf thymus DNA as primers (8-12 nucleotides long, a gift from C. Rice). The mass ratio of the RNA to primers is 1:2. After the synthesis of the first strand cDNA, the cDNA-mRNA hybrid was used as a substrate for treatments with RNase H, *E. coli* DNA polymerase I and *E. coli* DNA ligase. The detailed procedure for making double-strand cDNA was described by Watson and Jackson (3). The following EcoRI linker ligation, size fractionation and insertion of the double-strand DNA into the λgt10 vector was described previously (4).

Synthetic Nucleotides

Synthetic oligonucleotides were synthesized on an automatic DNA synthesizer using solid-phase phosphoramidite chemistry as described elsewhere (5-6). The oligomers were separated from smaller incomplete reaction products by 20% preparative polyacrylamide gel electrophoresis.

Screening λgt10 Testis cDNA Library with the Oligonucleotide Probe

Purified oligonucleotides were labeled using γ-^{32}P-ATP (3,000 Ci/mmole) and T4 polynucleotide kinase to the specific activity of 3×10^8 cmp/μg (7). *E. coli* C600 Δ Hfl cells were infected by the phage, plated on NZY plates and blotted to nitrocellulose filters as described somewhere else (4). Prehybridization was in 5X SSC 5X Denhardt's solution (8) and 100 μg/ml sheared denatured calf thymus DNA at 42° C for 2 hrs. The labeled oligonucleotides were then added to 3×10^6 cpm/ml and the hybridization was at 42°C for 16 hrs. The filters were washed with 5X SSC 0.1% SDS and 2X SSC 0.1% SDS solution at 42°C, dried and autoradiographed. 1X SSC is 0.15 M sodium chloride 0.015 M sodium citrate pH 7.

Genomic DNA and RNA Blot Analysis

Genomic DNA was prepared by the method of Blin and Stafford (9, made by J. Minor). The genomic blot was carried out as previously described (10). The prehybridization was in 5X SSC, 5X Denhardt's solution (8) and 100 μg/ml sheared denatured calf thymus DNA at 68°C for 2 hrs. The labeled 438 bp SalI/SalI fragment and 799 bp SalI/RI fragment (see Fig. 1a) were added to 10^6 cpm/ml and the hybridization was carried out for 16 hrs. at 68°C. The filters were washed in 5X SSC, 0.1% SDS for 1 hr, then in 2X SSC, 0.1% SDS at 68°C for 1 hr. Air-dried filters were autoradiographed.

A denaturing formaldehyde-agarose gel was used for RNA separation. It was described by Scheller *et al.* (11). The hybridization was in 5X SET, 1X Denhardt's solution (8), 0.025 M phosphate buffer (pH 6.8), 50% formamide, 50 µg/ml sheared and denatured calf thymus DNA and 10^6 cpm/ml nick-translated bindin cDNA probe. Filters were washed for 1 hr in 5X SET, 0.1% SDS and 1 hr in 2X SET, 0.1% SDS at 42°C.

Primer Extension Experiment

The primer was made and kinase labeled as described above. Two ng of primer was mixed with 1 µg testes $[pA]^+$ RNA or 16 µg of testes total RNA in 5 µl, 2mM phosphate buffer pH 6.8, heat denatured at 95°C for 3 minutes, and hybridized at 43°C for 1 hr in a total volume of 10 µl of 100 mM Tris-HCl buffer pH 8.6, 10 mM $MgCl_2$, 20 mM β-Mercaptoethanol, 1 mM phosphate buffer pH 6.8 and 140 mM KCl. Forty µl of this buffer containing in addition, 500 µM each deoxynucleotide triphosphate, 2 µg actinomycin D (Sigma, prepared in 80% ETOH at 0.8 mg/ml), and 30 units RNasin (Promega) was added and incubated with 25 units AMV reverse transcriptase (Life Science) at 43°C for 1 hr. Extension products were precipitated and run over a 40 cm 8% acrylamide urea denaturing gel.

References

1. J. M. Chirgwin, A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. Biochemistry 18, 5294 (1979).
2. H. Aviv and P. Leder. Proc. Natl. Acad. Sci. USA 69, 1408 (1972).
3. C. J. Watson and J. F. Jackson. In "DNA Cloning: A Practical Approach", Vol. 1, D. Glover, ed., (IRL Press, Oxford, 1985), p. 79.
4. T. V. Huynh, R. A. Young and R. W. Davis. In "DNA Cloning: A Practical Approach", Vol. 1. D. Glover, ed., (IRL Press, Oxford, 1985), p. 49.
5. S. L. Beaucage and M. H. Carruthers. Tet. Lett. 22, 1859 (1981).
6. M. D. Matteucci and M. H. Carruthers. Tet. Lett. 21, 719 (1980).
7. M. Houghton, A. G. Stewart, S. M. Doel, J. S. Emtage, M. A. W. Eaton, J. C. Smith, T. P. Patel, H. M. Lewis, A. G. Porter, J. R. Birch, T. Cartwright, and N. H. Carey. Nucl. Acid. Res. 8, 1913 (1980).
8. D. T. Denhardt. Biochem. Biophys. Res. Commun. 23, 641 (1966).
9. N. Blin and D. W. Stafford. Nucl. Acid Res. 3, 2303 (1976).
10. E. M. Southern. J. Mol. Biol. 98, 503 (1975).
11. R. H. Scheller, L. B. McAllister, W. R. Crain, D. S. Durica, J. W. Posakony, T. L. Thomas, R. J. Britten and E. H. Davidson. Mol. Cell. Biol. 1, 609 (1981).

CHAPTER 2

The Partial Characterization of the Bindin Gene

Introduction

The isolation and sequencing of bindin cDNA were presented in Chapter I. As shown in that chapter, bindin cDNA was used as a probe to study the expression of bindin in various tissues. It was demonstrated clearly that bindin gene is only expressed in testes but not in other tissues we tested (Chapter I, Fig. 4). It is obvious that bindin gene expression is spatially and developmentally regulated. It is important to know the bindin gene structure in order to study the mechanism of its regulation. The isolation and partial characterization of the bindin gene will therefore be described in this chapter.

Another problem will also be addressed in this chapter: whether the bindin gene exists in female sea urchins. As we discussed before, bindin gene expression is strictly male specific. Other studies have demonstrated that there are sperm-specific histone H1 and H2A genes in the sea urchin (1-3). On the other hand, the vitellogenin gene, a female-specific gene in a variety of organisms, including vertebrates, nematodes and insects (4,5), has been shown to be expressed in both female and male tissues in the sea urchin (6). The fact that only male-specific genes have been found has raised a fundamental question of sea urchin genetics, i.e., what is the sea urchin sex chromosome pattern? The study of sea urchin chromosomes has been a difficult task due to their small size and indistinguishable shape. Recently, however, Gerhart (7,8) did elegant studies on this subject in which he analyzed the chromosomes including the number per cell, the standard karyotypes and the response to routine chromosme banding techniques of six North American sea urchins. He did not observe chromosome sexual dimorphism in any of the six species he studied including *S. purpuratus*. Gerhart's observation cast serious doubts upon the earlier reports which concluded that the male sea urchin is the digametic sex (29). We thought bindin might be a good candidate to study the sea urchin sex chromosome pattern since bindin cDNA can be used to test whether

female sea urchins have bindin genes. If bindin genes were not found in female chromosomes, it would indicate that the male sea urchin has a specific chromosome. The human and mouse H-Y antigen is an example of a gene product encoded by Y chromosome DNA (9). On the other hand, if the females do have bindin genes, the question remains to be answered. The result of the genome blot in Chapter I shows that there are two bindin alleles (in the sense of the restriction enzyme polymorphism) in male sea urchin individuals. While it could be true that one of the alleles is of paternal origin and the other is of maternal origin, one cannot exclude the possibility that there are two copies of the bindin gene on the male chromosome. An experiment designed to answer this question was carried out and is presented in this chapter.

Results

The isolation and partial sequencing of bindin genomic clones.

The bindin cDNA described in the previous chapter was used as a probe to screen three sea urchin genomic libraries. The DNAs used for the construction of the three libraries were isolated from sperm of three individual sea urchins (*S. purpuratus*). The bindin cDNA was nick translated to a specific activity of 10^8 cpm/μg and the screening was performed at relatively high stringency [0.75 M Na^+ 68°C, (12,13) see Experimental Procedure]. Twenty equivalent genomic size of clones from each library were screened and only one positive has been identified which was named bindin gene clone F. The reason for the scarcity of the bindin clone in these libraries is unknown. It could be due to the presence of some "poison" sequence in bindin gene which slows down the growth of the infected cells or due to the recombination and the subsequent deletion of the DNA sequence around the bindin gene. Clone F DNA was isolated and analyzed. As shown in Fig. 1A, this clone contains a 14 kb insert DNA, and a 2.1 kb EcoRI

fragment of this insert hybridized to the 3' half of the bindin cDNA (shown on Fig. 1B and C). F does not hybridize to the middle portion (the 439 bp SalI/SalI fragment) of the bindin cDNA (shown on Fig. 1C). The 2.1 K EcoRI fragment also has restriction enzyme recognition sites corresponding to those of the 3' end of the bindin cDNA. Apparently the divergence occurs upstream of the SphI site at base pair 1542 of the bindin cDNA. This suggests that there is an intervening sequence in this area. Further sequencing of this region is necessary to determine the precise position of the intervening sequence.

In order to isolate other parts of the bindin gene, a λgt10 minigenomic library was made (see Experimental Procedures for detail). As shown in Fig. 2, a 1.3 kb DNA fragment from EcoRI complete digestion of three *S. purpuratus* individual DNAs (provided by J. L. Minor in E. H. Davidson's lab) hybridized with bindin cDNA. (This procedure was described in Chapter I.) DNA of individual #2 was completely digested with EcoRI and was separated on a preparative agarose gel. The DNA of around 1.3 kb was electroeluted and was ligated to λgt10 arms. The ligated DNA was used to infect the *E. coli* C600 ΔHfl strain and the recombinant phage was screened by nick-translated bindin cDNA (10^8 cpm/μg) (12). One of the clones, gt10I, hybridized to the probe was shown to have a 1.3 kb insert. The restriction map, the sequencing strategy and the partial sequence of this DNA are shown in Fig. 3 (14,15). This clone has a 219 bp middle region identical to bindin cDNA from base pair 124 to base pair 344 (boxed sequence in Fig. 3). This region turned out to be an exon and the two flanking fragments to be introns for the reasons given below. I will refer the intron 5' of the exon as intron I and the intron 3' of the exon as intron II for convenience.

Intron I contains 579 bp and was sequenced completely. It has dinucleotide AG at the 3' boundary to the exon, which obeys the AG rule as the intron/exon junction sequence. Intron I is also pyrimidine rich at the 3' end with the sequence

Fig. 1A. *The restriction map of the F genomic clone.*

The filled bar indicates the 2.1 K EcoRI fragment of F genomic clone which hybridize to bindin cDNA. The open bars indicate the fragments of clone F derived from the EcoRI/XhoI digestion. The hatched bars indicate the cDNA fragments used as probes in Fig. 1C. The t#16 bindin cDNA contains the 3' half of the bindin cDNA and the downstream region which has not been sequenced.

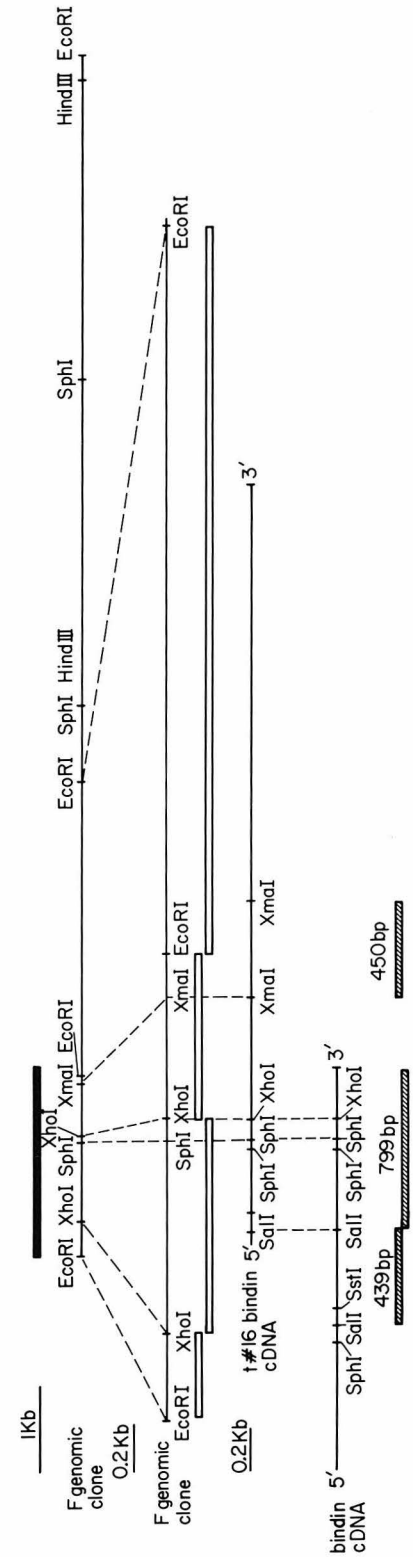

1Kb
F genomic clone
0.2Kb
F genomic clone
0.2Kb
bindin cDNA
t#16 bindin cDNA
EcoRI
XhoI
SphI
XmaI
HindIII
5'
3'
450bp
439bp
799bp
SalI
SstI

Fig. 1B. *The hybridization of F genomic clone with bindin cDNA.*

The F genomic clone DNA was digested with the restriction enzymes indicated and separated on a 1% agarose gel. The gel was denatured and blotted to the nitrocellular filter as the standard procedures (28). The probe was the nick-translated bindin cDNA. The arrow indicates the 2.1 K EcoRI fragment that hybridized to the probe.

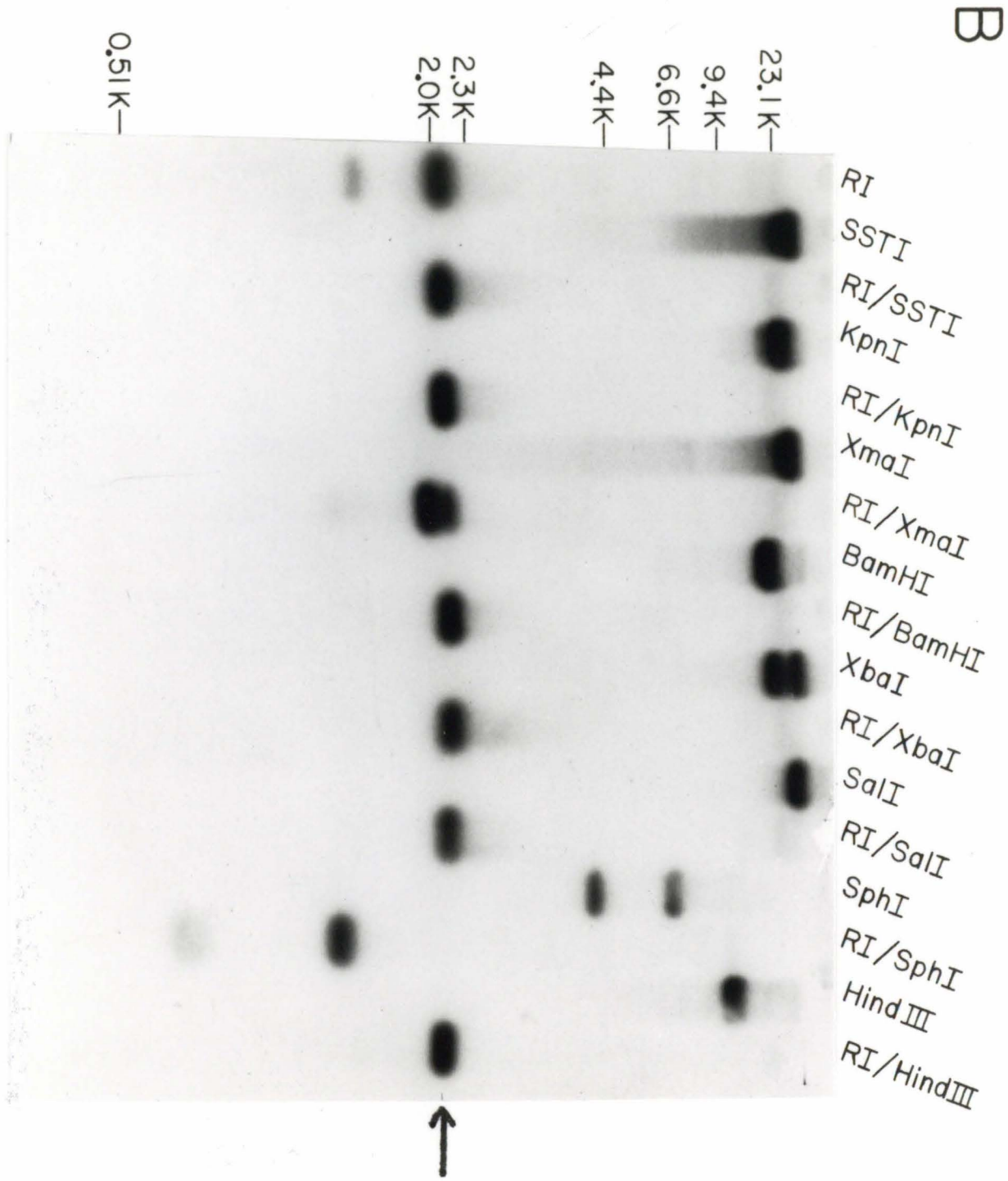

B
23.1K
9.4K
6.6K
4.4K
2.3K
2.0K
0.51K
RI
SSTI
RI/SSTI
KpnI
RI/KpnI
XmaI
RI/XmaI
BamHI
RI/BamHI
XbaI
RI/XbaI
SalI
RI/SalI
SphI
RI/SphI
HindIII
RI/HindIII

Fig. 1C. *The hybridization of EcoRI/XhoI digested F genomic clone with the bindin cDNA fragments.*

The F genomic DNA was digested with EcoRI/XhoI (lane 1.3, and the open bars in Fig. 1A) and the bindin cDNA was digested with SalI/EcoRI (lanes 2,4) and the DNA was separated on a 1% agarose gel. The gel was denatured and blotted to the nitrocellulose filter by the standard procedures (28). The probes were the nick-translated bindin cDNA fragment, 439 bp SalI/SalI fragment (the middle fragment of bindin cDNA) in lanes 1 and 2, and the 799 bp SalI/EcoRI fragment (the 3' half of the bindin cDNA) in lanes 3 and 4.

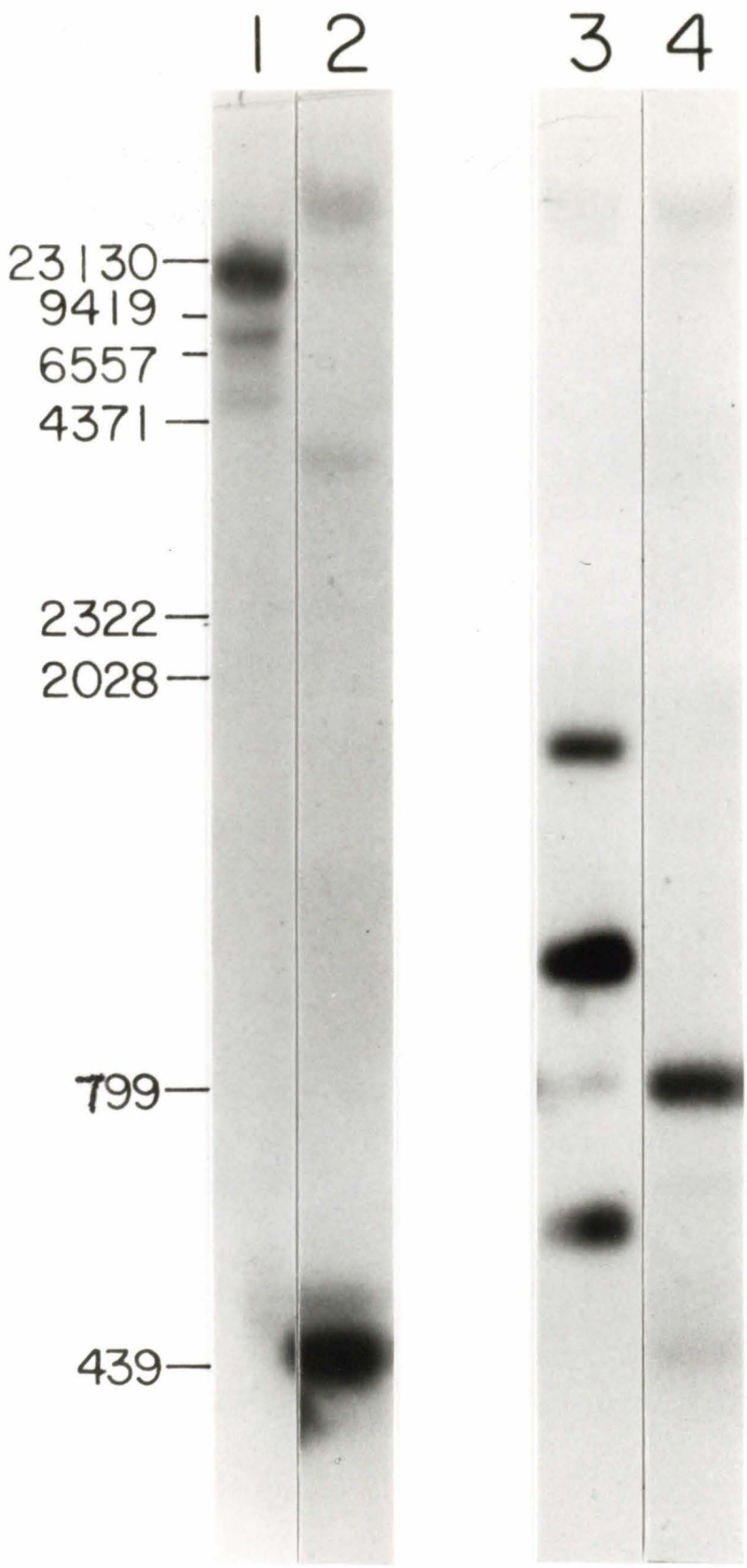

1
2
3
4
23130
9419
6557
4371
2322
2028
T99
439

Fig. 2. *The genomic blot of sea urchin individual sperm DNA.*

Sperm DNA was digested completely with Eco RI and was separated on a 1% agarose gel. The gel was denatured and blotted as the standard procedure (28). The probe was the nick-translated bindin cDNA. The sperm DNA used in each lane was from lane A, individual #2, B, individual #5, C, individual #11. The arrow indicates the band which was isolated for making λgt10 minigenomic library.

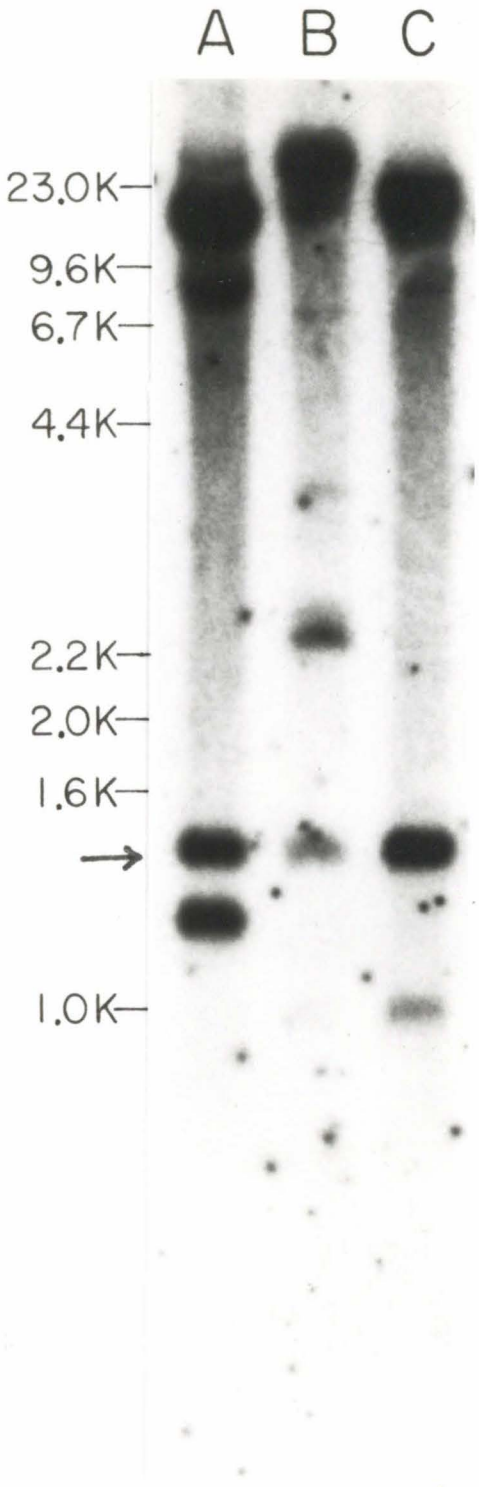

A
B
C
23.0K
9.6K
6.7K
4.4K
2.2K
2.0K
1.6K
1.0K

Fig. 3A. *The restriction map and the sequence strategy of gt10I.*

The filled bar indicate the bindin cDNA. The hatched bar indicates the gt10I fragment that is homologous with bindin cDNA. The open bars indicate the gt10I DNA fragments that are not homologous with bindin cDNA. The lengths of the arrows indicate the direction and number of nucleotides for which sequence was determined by the dideoxynucleotide chain termination method.

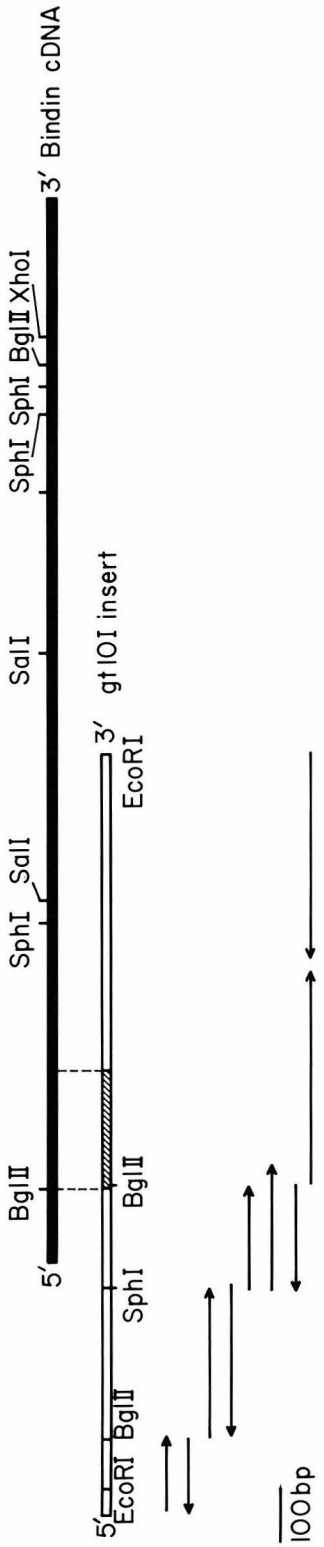

3′ Bindin cDNA
XhoI
BglII
SphI
SphI
SalI
gt10I insert
3′
EcoRI
SalI
SphI
BglII
BglII
5′
SphI
BglII
5′
EcoRI
100bp

Fig. 3B. *The partial sequence of gt10I.*

The nucleotides that were not determined are indicated by **n.** The lower cases indicate the ambiguous nucleotides. The underlined sequences indicate the G and C stretches. The sequence in the boxed area is homologous with bindin cDNA from base pairs 124 to 344.

Fig 3B


```
1     GAATTCGACT CAATTTTGCA CAACCTAAGA ATTAGAATAC GGTCAGTCAC CCCCCCCCCA
61    TGCGCGCCCA cTTTCACACG CGCATATCCC TCCCCCCCCC CCCCCAGTTT TTTCCCCTTG
121   TTCCGGCATG TGTCATTAGA TCTAGTGTCA TGTCCCTGGG CCCCGtAGCA TAAcAACTTA
181   CCACTGACTG GTAACTTAGT CTctGCTATC aaTagcCAAC TATCATGGTc AAgTaGGGTT
241   TCCATGGTAG TTACCTTTGA TGGCAAgAGT TACCCAgTCA AcTGGCAAGT TTcGTTGCAA
321   CGGGGTCCGT GTAAAGTCgA ATTTTCCtAT ATGTGATGCC GGACTTTGGT GCGTAAAAAG
381   TAACCCTTTA CCGGACATTT TGGTAACACA CGCATGCGGT CATTCCAACC CCGGAgTgCG
441   GGGGGGGGGG GGGGGaAGGC GGGAGgTATC CCCCACTCcT cACGGAGTCg ACGACAGACA
501   cggcTAATGA TCGAcTGTTT TCTAGACATA TcATGCGCAC GACGATGTGC AGTTGCTCAA
561   GATGATAATG ATGATAGTTC TTTGTTTTTA TTTTTATAGA TCTACATTTA GCATCATGGG
621   TTTCCATCAA ATTTTAGTCA CTGTAGTTGC CCTCGCTTTA GCCTCTGTCA GAGCCGAGTT
681   CCCATCCCGT ACCGACTCCC CTACTGACTG CCCCGAAGCA GATCAAGGTG CTGGTGTCGC
741   GGCTCCTTCG CTCAGTGCTG GAGAACcGTA TGAAGAGGCA GGAATGACAG GAGAAATTGG
801   AAACAGAATT ACAAAGCTGT AAGTTTAGAT TTGGCGTTTT CCATGAGATG ACAAAGTTCA
861   TCACATCATT TGAAATTGCG TGTACAGGAT CACAGCGATG TCCATGGAAG ATTGCACTAC
921   AGGTCTAGAT TGGCAGCTAC ACTTACTGGT GTAACGAGCA ATTTCGCCAT GCGAAGCTTA
981   CGGTTGGACT GGTCACnnnn nnnnnnnnnn nnnnnnnnnT CGTAACGTTC CAAAGCTCAA
1041   CGTTCATGAA AGCCTTACGG ACATCTCTGA TGCTAAACGG AAGCAAGATG CTTGCACACT
1101  TAAAGACATC AGAGTCTCAT TGACAAGGTC TTGACTTGGA AAGAAAGTAG AAGGTACCCA
1161  GACCAAGACG GGAAACCAAG AACTAGATTT GGGAAGGGGT GAATTGGAAG TGCCCACCAC
1201  TACGAAGAGT AGTACATGGG CGGGCAGAAA ACAGAGATGT GTCGAAAAAG TGTTGCAAGG
1261  AACCTCGTGA GCTAATAAAA GATACCTgAA TAGCGTAGGA ATGGGTATAT TCGAACATAA
1321  AAATAGAGGC AAATTtAAAT TATTTGTGTC TCTTTT
```

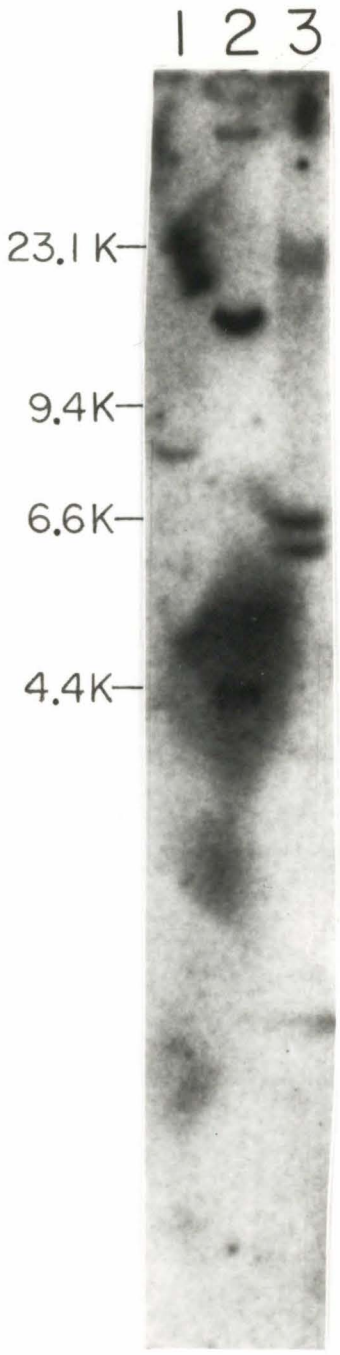
T5ATAG which resembles the acceptor intron consensus sequence $(\begin{smallmatrix}T\\C\end{smallmatrix})$ nN $\begin{smallmatrix}C\\T\end{smallmatrix}$AG (16). In addition, it has been shown that there is a sea urchin potential 3' splice signal A $\begin{smallmatrix}C\\T\end{smallmatrix}$TAAT 24 to 33 bp from the 3' splice point (17). There is an ATAAT sequence 30 bp from the 3' end of this fragment, which is close to the consensus sequence. Intron II was sequenced except for the middle 20 bp. The 5' sequence GTAAGT of this intron is identical to the consensus sequence of the intron donor sequence (16). The data shown above indicate that there are an exon and two introns lying in this 1.3 K genomic fragment. The introns are located 171 bp and 391 bp from the transcription initiation site respectively. Intron I is in the leader sequence of the bindin messenger RNA. The AUG translation initiation site lies 5 nucleotides into the exon (see Chapter I, Fig. 1). Although there could be an exon lying in the middle of the 20 bp unsequenced region of intron II, it is unlikely unless the exon is smaller than 20 bp. Further sequencing data are necessary to answer this question.

The study of the bindin gene in female sea urchins.

Female coelomocytes were chosen as the DNA source for this study for the following reasons. There are enough cells from one individual to perform the analysis and it is relatively easy to lyse the cells and purify the DNA. The method for isolating the DNA is described in the Experimental Procedures. Isolated DNA was subjected to genomic analysis. The probe was nick-translated bindin cDNA with specific activity of 10^8 cpm/μg. Male coelomocyte DNA was applied as a control for somatic cells (not shown). As shown in Fig. 4, one of the females has two restriction fragments hybridizing to the bindin probe, which probably represents the two alleles of bindin gene and the other female has one which indicates the homozygosity of the two bindin alleles in this female. Thus the bindin gene exists in female sea urchins although it is not expressed.

Fig. 4. *The genomic blot of coelomocyte DNA of female sea urchin.*

The DNA was digested with BglII and was separated on a 1% agarose gel. The gel was denatured and blotted as the standard procedure (28). The probe was the nick-translated bindin cDNA. The DNA used in each lane was isolated from (1) coelomocytes of female individual #1, (2) coelomocytes of female individual #2, (3) sperm of male individual #11.


1 2 3
23.1 K
9.4K
6.6K
4.4K

Discussion

The data presented above show that the bindin gene is discontinuous and is interrupted by at least three introns. It is interesting that other non-histone sea urchin genes which have been studied all share the common feature of an intron close to the transcription initiation site. For example, sea urchin spec I gene (*S. purpuratus* ectoderm specific gene; 18) has an intron 220 bp from the transcription initiation site and is right after the AUG translation initiation codon; there is an intron in the 5' non-translated region of CyI and CyIIIa cytoplasmic actins (19). In the case of the sea urchin major spicule matrix protein, the intron is located 215 bp from the transcription initiation site and 37 amino acids from the translation initiation codon (unpublished data from H. M. Sucov in E. H. Davidson's lab). In other systems where the introns close to the promoter sequence bear enhancer or regulatory sequences, for example, a tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene (20,21). The first intron of the human growth hormone gene contains binding sites for glucocorticoid receptors (22). The significance of these introns in sea urchins is unknown.

It should be noted that the first intron fragment in λgt10I has two stretches of CS, C10, and C14, separately, and it also has a row of 16 Gs (the underlined sequences in Fig. 3). These C and G tracts could form stable stem and loop structures in the bindin RNA precursor according to the thermodynamic calculations (23). The conformation of the bindin nuclear precursor RNA in the nucleus is unknown. Whether it could serve any function such as facilitating the RNA processing or regulating gene activity is an open question. A controlled deletion and transformation experiment would be the direct way to study the function of these sequences.

Experimental Procedures

The construction of genomic libraries

The name and the description of the libraries are as follows: "Cyril" genomic library was made from DNA partially digested with HaeIII, ligated with EcoRI linkers and inserted into λ Charon 4A arms (10, made by M. Chamberlin in E. H. Davidson's lab); "Fred" library was made from EcoRI partially digested DNA ligated with λ Charon 4A arms (10, made by H. T. Jacobs in E. H. Davidson's lab); "Bimbo" library was made with MboII partially digested DNA, ligated with BamHI linkers, and inserted into BamHI digested EMBL3 λ vector (11, made by F. J. Calzone in E. H. Davidson's lab).

The construction of the λgt10 minigenomic library.

50 μg of individual sea urchin #2 DNA made by J. J. Minor according to the method described elsewhere (24), was completely digested by restriction enzyme EcoRI and separated on a 1% preparative agarose gel. The agarose gel region containing 1.3 kb DNA was cut out. The DNA was eluted from the gel slice by electroelution in dialysis tubing (VWR Scientific, average pore radius permeability of 24 Å) in buffer A [0.04 M Tris-acetate, 0.002 M EDTA (25)] at 75 volts for 100 minutes. The DNA solution was then loaded on a "Elutip" column (Schleicher & Schnell) and was eluted by a high salt solution containing 1 M NaCl, 20 mM Tris HCl (pH 7.4), 1 mM EDTA. ETOH was added to precipitate the DNA. The DNA was recovered by spinning in a microfuge for 15 minutes at 4°C. The DNA pellet was resuspended in 5 λ of TE buffer (10 mM Tris HCl pH 8, 1 mM EDTA). 0.5 μl of this DNA and 1 μg of λgt10 arms (purchased from Vector Cloning System) were ligated in 5 μl of ligation buffer containing 66 mM Tris HCl pH 7.5, 5 mM $MgCl_2$, 5 mM dithiothreitol, 1 mM ATP. 1 U of T4 DNA ligase (purchased from BRL) was added and the reaction incubated at 4°C for 10 hours. The ligated DNA was

packaged in a packaging extract [made by J. L. Minor in E. H. Davidson's lab (26)]. *E. coli* 600 ΔHfl cells were infected and NZY plates were used to plate the phage (27). There are 2.5×10^5 recombinants in this library.

The isolation of female sea urchin coelomocyte DNA.

10 ml of 0.1 M EGTA, MPFSW (Millipore filtered sea water) was added to a 50 ml centrifuge tube. The tube was topped by a funnel, covered by cheesecloth soaked with the same solution. A female sea urchin was dissected by cutting out the lantern tissue. The coelomic fluid was poured into the funnel (there are about 7-15 ml of the coelomic fluid in each sea urchin). The tube was spun in a tabletop centrifuge for 1 minute at the maximal speed. About 0.2 ml of cells were precipitated on the bottom of the tube. An equal volume of 2X lysis buffer and protease K was added. (The final concentration is 50 mM KCl, 200 mM EDTA, 50 mM Tris HCl, pH 7.4, 1.2% sarcosine and 100 μg/ml of protease K). The reaction was carried out at 55°C for 3 hours. The lysate was phenol-chloroform extracted twice and chloroform extracted once. LiCl was added to the aqueous phase to the final concentration of 0.3 M. The DNA was precipitated by ETOH and was recovered by spinning as described above. The DNA pellet was resuspended in TE buffer. Half of the DNA was subject to the further restriction enzyme digestion and was loaded on one lane for the genomic blot analysis.

DNA sequence analysis

DNA sequences were determined by the method of dideoxy nucleotide chain termination (14,15).

Acknowledgements

I thank J. A. Ferrier for helping me mapping the F clone and isolating the 1.3 k DNA for the minigenomic library. I also thank J. J. Minor for providing me with the individual sea urchin genomic DNA and the *in vitro* packaging extract.

References

1. Strickland, M., Strickland, W. N., Brandt, W. F., and Von Holt, C. (1977). *Eur. J. Biochem.* **77**, 263.
2. Strickland, W. N., M. Strickland, de Groot, P. C., Von Holt, C., and Wittman-Liebold, B. (1980). *Eur. J. Biochem.* **104**, 559.
3. Busslinger, M. C. and Barberis, A. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 5676.
4. Kimble, J. and Sharrock, W. J. (1983). *Dev. Biol.* **96**, 189.
5. Brennan, M. D., Weiner, A. J., Goralski, T. J., and Mahowald, A. P. (1982). *Dev. Biol.* **89**, 225.
6. Shyu, A., Raff, R. A., and Blumenthal, T. (1986). *Proc. Natl. Acad. Sci. USA*, in press.
7. Gerhart, S. G., Wagenaar, E. B., and Lin, C. C. (1979). *Genetics* **91**, 538.
8. Gerhart, S. G. (1983). Ph.D. Thesis from the University of Calgary.
9. Ohno, S. (1976). *Cell* **7**, 315.
10. Maniatis, T., Fritch, E. F., and Sambrook, J. (1982). In *Molecular Cloning*. Cold Spring Harbor Laboratory, pp. 31 and 269.
11. Frischauf, A., Lehrach, H., Pouska, A., and Murry, N. (1983). *J. Mol. Biol.* **170**, 827.
12. Rigby, P. W. J., Dieckmann, M., Rhodes, C., and Berg, P. (1977). *J. Mol. Biol.* **113**, 237.
13. Maniatis, T., Fritch, E. F., and Sambrook, J. (1982). In *Molecular Cloning*. Cold Spring Harbor Laboratory, p. 309.
14. Sanger, F., Nicklen, S., and Coulson, A. R. (1977). *Proc. Natl. Acad. Sci. USA* **74**, 5463.
15. Messing, J. (1983). *Meth. Enzymol.* **101**, 20.
16. Mount, S. M. (1982). *Nucl. Acids Res.* **10**, 459.

17. Keller, E. B. and Noon, W. A. (1984). *Proc. Natl. Acad. Sci. USA* **81**, 7417.

18. Houck, S. L., Carpenter, C. D., Hardin, P. E., Bruskin, A. M., and Klein, W. H. (1986). *J. Mol. Biol.*, in press.

19. Schuler, M. A., McOsker, P., and Keller, E. B. (1983). *Mol. Cell Biol.* **3**, 448.

20. Gillies, S. D., Morrison, S. L., Oi, V. T., and Tonegawa, S. (1983). *Cell* **33**, 717.

21. Banerji, J., Olson, L., and Schaffner, W. (1983). *Cell* **33**, 729.

22. Moore, D. D., Marks, A. R., Buckley, D. I., Kapler, G., Payvar, F., and Goodman, H. M. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 699.

23. Boyer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974). *J. Mol. Biol.* **86**, 843.

24. Blin, N. and Stafford, D. W. (1976). *Nucl. Acids Res.* **3**, 2303.

25. Maniatis, T., Fritch, E. F., and Sambrook, I. (1982). In *Molecular Cloning.* Cold Spring Harbor Laboratory, p. 156.

26. Mullins, J. I., Casey, J. W., Nicolson, M. O., Burck, K. B., and Davidson, N. (1981). *J. Virol.* **81**, 688.

27. Huynh, T. V., Young, R. A., and Davis, R. W. (1985). In *DNA Cloning. A Practical Approach.* (D. Glover, ed.). JRL Press, Oxford.

28. Southern, E. M. (1975). *J. Mol. Biol.* **98**, 503.

29. Harvey, E. B. (1956). *Pubbl. Staz. Zool. Napoli* **28**, 127.

CHAPTER 3

Summary and Discussion

A study of the bindin mRNA, the bindin gene, and its expression are presented in this thesis. The structural analysis explores several interesting features of the bindin protein, which I have already discussed in the previous chapters. I would like to review and speculate on the possible functions of bindin based on some recent works done by others.

The Questions Concerning the Bindin Precursor

The sequence of the bindin cDNA clone shows that it encodes a larger protein precursor with the mature bindin located at the carboxy terminus of this molecule (shown in Fig. 1B, Chapter I). The immunoprecipitation experiment (Fig. 3 Chapter I) further supports the existence of the larger precursor. The four basic amino acids right in front of the N terminal of the mature bindin protein suggests that the mature bindin is the cleavage product of proteolytic enzymes. One question immediately raised by this result is the function of the precursor, if any. Many proteins including cellular (1-4), viral (5-8), and organelle (9) proteins are made by larger precursors. The known functions for precursors include facilitating the intracellular transport of the protein (9), ensuring the correct folding of the mature protein (1,10), and keeping the proteins in inactive forms prior to utilization (5-8,11-14). The processing of bindin protein may resemble the processing of collagen (discussed in Chapter I). Another example of precursor processing is the 65 kd fusion protein (F protein) of Sendai virus (5) which is neither fusogenic nor infective in the cell unless subjected to mild trypsin digestion which results in an active 51 kd protein. It has been shown that the activation of the F protein involves a conformation change with exposure of a new hydrophobic region which is located at the N terminal of the mature F polypeptides (5). The precursor form provides a means of preventing the F protein from engaging in hydrophobic interaction with intracellular membranes before it

is assembled into the viral envelope. The processing of bindin has parallels with the maturation of F protein since bindin has been shown to fuse phospholipid vesicles *in vitro* (15, which will be discussed below). The first 56 amino acids in the N terminal region of mature bindin contain many non-polar amino acids and no charged amino acids, which have been suggested to interact with membrane hydrophobically (16). It might be that the cleavage of the bindin precursor also results in the exposure of this N terminal hydrophobic region and thereafter the activation of mature bindin protein. The mitochondrial ATP synthase is another example of a protein processed by cleavage. It is synthesized in the cytoplasm from a nuclear-encoded message as a 147 amino acid soluble preprotein, from which a 66 amino acid N terminal polypeptide believed to be required for transport into the mitochondria is subsequently cleaved off (9). The same strategy could apply for bindin, that it may be synthesized in the rough endoplasmic reticulum of the spermatocytes, transported to the Golgi complex, and later incorporated in the membrane-bound acrosome granule (40). The N-terminal polypeptide could include recognition sequences that function during this process in targeting the molecule to its ultimate location. Another possibility is that the pro-polypeptide of bindin is itself a functional protein. Bindin protein sequence was compared with the National Biomedical Foundation protein sequence database, and no proteins share a significant degree of homology with the bindin precursor. However, enzyme activities including phospholipase (17), chymotrypsin (18) and trypsin-like protease (19) are released during the acrosome reaction, in which the pre-polypeptide of bindin may participate. No attempt has been made to detect the pre-bindin polypeptide since it was not known until the present work. Although mature bindin is the major component of the acrosome granule, the propeptide may be present, but not be detected. It could be solubilized by the isolation buffer or associated with the membrane of the acrosome granule and is

separated from the acrosome granule during the preparation of mature bindin proteins. The suggested functions of the bindin precursor discussed above may be tested or ruled out experimentally. An antibody could be raised specifically against the pro-polypeptide and the antigen tested for localization in the sperm before and after the acrosome reaction. If the pro-polypeptide is not found in the intact acrosome then this suggests that it has been removed prior to bindin accumulation in the acrosome and therefore may also function to transport the mature protein. If it is found at the egg-sperm junction just as the mature bindin is, this indicates that the propeptide is not degraded after proteolytic cleavage and is therefore probably a functional protein. An *in vitro* pulse labeling experiment (by dissecting testes tissue and incubating it with ^{35}S-methionine sea water and chasing with cold methionine) and the subsequent protein gel analysis could indicate the synthesis and the degradation of the bindin precursor if it is abundant enough to be detected. A Western blot analysis can be done by isolating the total sperm protein and probing with antibodies specifically against the pro-polypeptides and the size of the reacted band can suggest the fate of the pre-polypeptide.

The Mechanism of Sperm-Egg Adhesion

There are several lines of evidence which indicate that the mature bindin is the species-specific mediator of sperm-egg adhesion. First, bindin has been shown to be at the point of sperm-egg attachment, and is exposed only in acrosome-reacted sperm (20). Second, isolated particulate bindin agglutinates intact unfertilized eggs in a species-specific fashion (21,22). In order to perform this binding function, bindin must interact with both the vitelline layer of eggs and the acrosome process. The sulfated glycoconjugates of the egg vitelline layer have been suggested to be the bindin receptor (23,24). On the other hand, the

interaction and association of bindin with the acrosomal process is not clear. The protein sequence of bindin indicates that it is unlikely to be an integral membrane protein since it does not contain a typical hydrophobic membrane spanning region (25). This agrees well with the fact that bindin is isolated from the demembranated acrosome granule. However, there are two regions with the potential to interact with the phospholipid bilayer of the acrosome process hydrophobically. One region previously suggested by Vacquier and Moy (16) is the 56 amino acid N-terminal of the mature bindin which does not contain any charged amino acids. The other region is from amino acid 82 to 95 of the mature bindin containing 9 hydrophobic amino acids in this area (shown under the bracket in fig. 3, Chapter I). From the primary structure analysis, these are the two regions likely to be associated with acrosome process membrane, although some hydrophobic regions derived from the secondary structure may be involved with the interaction. Recently, Glabe (26) designed an *in vitro* system to test the interaction of bindin with phospholipid vesicles, and found that bindin is specifically associated with phospholipid bilayers in a gel state. Both gel phase and liquid phase domains have been postulated in the membrane of the acrosome process (27,28). The experiments show that the type and the charge of the head group of the phospholipid bilayer play a relatively minor role in the interaction, suggesting instead that the hydrophobic interaction of bindin and the hydrocarbon region of the fatty acid moieties of the lipid bilayer are responsible for the association of bindin with the phospholipid vesicles. Glabe further demonstrated that bindin associated with the bilayer is still sensitive to cleavage by trypsin, which suggests that most of the polypeptide chain remains exposed at the surface of the membrane. This system may be used for future study by altering the bindin protein in specific areas via gene manipulation and localizing the sequences which are responsible for the association of bindin with membrane bilayers.

The Possible Fusogenic Function of Bindin

Bindin is present at the site of the attachment of sperm and egg, which is the same place that the membrane of the acrosome process fuses with the egg plasma membrane (25). When investigating sperm agglutination induced by egg jelly, Collins (29) observed in a few cases that the acrosomal process of one sperm adhered to the midpiece of another by fusion of the acrosomal tubule and midpiece plasma membranes. In the transmission electron microscope Collins demonstrated clearly that the extracellular material derived from the exocytosis of the acrosome granule covers the junction of the acrosome tubule and the plasma membrane. Although bindin had not been identified at that time, the extracellular material he showed was subsequently identified as bindin (20) since it resembles bindin in all respects. This observation explores the possible involvement of bindin in sperm-egg fusion. The fusogenicity of bindin was directly tested by Glabe (30), who had already shown that bindin can associate with phospholipid bilayers in a gel state (26). Glabe incubated vesicles containing phosphatidylcholine and phosphatidylserine and showed that fusion occurs in the presence of bindin protein. An electron micrograph of negatively stained vesicles in the reaction revealed that the contact between two adjacent vesicles is apparently flattened, which may increase the contact area between the two vesicles and results in the fusion. This interaction is not observed in the absence of bindin. Based on these observations Glabe suggested that bindin may play a dual role in fertilization in that it recognizes and associates with bindin receptor on the egg vitelline layer and also participates in the actual fusion of the sperm and egg plasma membrane. It may be that bindin is a fusogenic protein that the specific interaction of bindin with bindin receptor enhances this function and makes the non-specific fusion less significant.

Bindin and the Reproductive Isolation

In nature the two closely related sympatric sea urchins, *S. purpuratus* and *S. franciscanus*, inhabit the same geographical area (often can be seen on the same rock), and they have overlapped spawning seasons. Hybrids between the two species can be found although very rarely (33). Artificial insemination between the gametes of the two species yields very low percentage of fertilization under conditions such that a very high percentage of fertilization occurs between the homologous gametes and the gametes from sea urchin *Lytechinus pictus* and *S. purpuratus* (34). However, a few viable hybrid embryos between *S. purpuratus* and *S. francsicanus* do develop, and these hybrid embryos have been a favored material for studying maternal and paternal inheritence in embryogenesis (34,35). Thus, the interspecific fertilization barrier, rather than developmental incompatibility, is probably the limiting factor which prevents interbreeding between the two species. There are several events occurring in the fertilization process, including the acrosome reaction, the adhesion and binding of sperm to egg, the fusion of the two plasma membranes, and the fusion of the two nuclei. It has been shown that both the induction of the acrosome reaction and sperm adhesion contribute to the overall specificity of fertilization, although the extent of the species specificity in these two events may vary between species (36-38). Brandriff and Vacquier (39) have shown that the acrosome reaction of *S. purpuratus* sperm can be induced by the jelly coat isolated from *S. franciscanus* and vice versa, which suggests that the induction of the acrosome reaction is not the barrier to cross fertilization between these particular species, though it is for other species. Therefore the adhesion and the binding of the gametes may play important roles in successful fertilization. Thus bindin, the protein mediating adhesion and binding, is probably the critical molecule for species specific

fertilization in these two species. The bindin-mediated egg agglutination experiments further support this hypothesis, in which eggs of *S. purpuratus* and *S. franciscanus* are agglutinated only by homologous bindin. Bindin isolated from *S. purpuratus* does not agglutinate *S. franciscanus* eggs and vice versa (21). The above experiments are consistent with the area that the interaction of bindin and its receptor (discussed in Chapter 1), is the main mechanism for the reproductive isolation of the two species. This hypothesis may be tested experimentally. One could introduce the bindin gene of *S. purpuratus* into *S. franciscanus* eggs, fertilize these eggs with *S. franciscanus* sperm, raise the fertilized eggs to the adult, and test the sperm of these adults in fertilizing eggs of *S. purpuratus* (assuming the transformation results in the stable chromosome integration and the expression of the gene is regulated as the endogenous bindin gene). It would be expected that a much higher interspecies fertilization rate would be obtained compared with the control if the above hypothesis is correct. If the experiment does support the function of bindin in reproductive isolation in these two species, one may further ask if bindin was involved in the evolutionary process leading to the formation of the two species since reproductive isolation is a requisite for speciation. The answer to this question should certainly extend our understanding of evolution and speciation.

References

1. P. Bornstein (1980). *Ann. Rev. Biochem.* **49**, 957.
2. J. H. Fessler, L. J. Fessler (1978). *Ann. Rev. Biochem.* **47**, 129.
3. H. S. Wiley, R. A. Wallace (1981). *J. Biol. Chem.* **256**, 8626.
4. G. Kreil, C. Molloy, R. Kachnitz, L. Haiaml, U. Vilas (1980). *Ann. N. Y. Acad. Sci.* **343**, 338.
5. M. -J. Gething, J. White and M. Waterfield (1980). *Proc. Natl. Acad. Sci. USA* **75**, 2737.
6. C. M. Rice, E. M. Lenches, S. R. Eddy, S. J. Shin, R. L. Sheets and J. H. Strauss (1985). *Science* **229**, 726.
7. J. Blok, G. M. Air, W. G. Laver, C. W. Wood, G. G. Lilley, E. F. Woods, C. M. Roxburgh and A. S. Inglis (1982). *Virology* **119**, 109.
8. P. L. Collins, Y. T. Hung and G. W. Wertz (1984). *Proc. Natl. Acad. Sci USA* **81**, 7683.
9. A. Viebrock, A. Perz and W. Sebald (1982). *The EMBO J.* **5**, 565.
10. R. E. Chance, R. M. Ellis and W. W. Bromer (1968). *Science* **161**, 165.
11. B. Kassell and J. Kay (1973). *Science* **180**, 1022.
12. H. T. Wright (1973). *J. Mol. Biol.* **79**, 13.
13. O. D. Ratnoff in *The Metabolic Basis of Inherited Disease* (3rd edition). J. B. Stanbury, J. B. Wyngaarden, D. S. Frederickson, eds. (McGraw-Hill, 1974) p. 1671.
14. O. D. Ratnoff and B. Bennett (1973). *Science* **179**, 1291.
15. C. G. Glabe (1985). *J. Cell Biol.* **100**, 800.
16. V. D. Vacquier and G. W. Moy in *Cell Reproduction*, E. R. Dirksen, D. m. Prescott and C. F. Fox, eds. (Academic Press, 1978), p. 379.
17. C. B. Metz (1978). *Curr. Top. Devel. Biol.* **12**, 107.
18. J. G. Csernansky, M. Zimmerman and W. Troll (1979). *Devel. Biol.* **70**, 283.

19. Y. Yameda and K. Aketa (1982). *Develop. Growth and Differ.* **24(2),** 125.

20. G. W. Moy and V. D. Vacquier (1979). *Current Top. Devel. Biol.* **13,** 31.

21. C. G. Glabe and V. D. Vacquier (1977). *Nature (Lond.)* **267,** 836.

22. C. G. Glabe and W. J. Lennarz (1979). *J. Cell Biol.* **83,** 595.

23. C. G. Glabe and V. D. Vacquier (1978). *Proc. Natl. Acad. Sci USA* **75,** 881.

24. D. P. Rossignol, B. J. Earles, G. L. Decker and W. J. Lennarz (1984). *Develop. Biol.* **104,** 308.

25. G. A. Adams and J. K. Rose (1985). *Cell* **41,** 1007.

26. C. G. Glabe (1985). *J. Cell Biol.* **100,** 794.

27. E. L. Bearer and D. S. Friend (1982). *J. Cell. Biol.* **92,** 604.

28. D. E. Wolf, W. H. Kinsey, and W. J. Lennarz (1981). *Devel. Biol.* **8,** 133.

29. F. Collins (1976). *Develop. Biol.* **49,** 381.

30. C. G. Glabe (1985). *J. Cell Biol.* **100,** 800.

31. A. Fridberger, J. Sundelin, V. D. Vacquier and P. A. Peterson (1985). *J. Biol. Chem.* **260,** 9092.

32. L. A. Jaffe (1976). *Nature* **261,** 68.

33. E. F. Swan (1953). *Evolution* **7,** 269.

34. R. R. Chaffee and D. Maza (1963). *Dev. Biol.* **7,** 502.

35. A. R. Moore (1943). *J. Exptl. Zool.* **94,** 211.

36. R. G. Summers and B. L. Hylander (1975). *Exp. Cell Res.* **96,** 63.

37. G. K. Segall and W. J. Lennarz (1979). *Dev. Biol.* **71,** 33.

38. C. Glabe, M. Buchalter, and W. J. Lennarz (1981). *Dev. Biol.* **84,** 397.

39. V. D. Vacquier (1979). *Amer. Zool.* 839.

40. M. H. Burgos and D. W. Fawcett (1956). *J. Biophysic. Biochem. Cytol.* **2,** 223.