The Journal of Immunology



# Sea urchin genes expressed in activated coelomocytes are identified by expressed sequence tags. Complement homologues and other putative immune response genes suggest immune system homology within the deuterostomes.

L C Smith, L Chang, R J Britten and E H Davidson

*J Immunol* 1996; 156:593-602; ;
http://www.jimmunol.org/content/156/2/593



*The Journal of Immunology* is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

# Sea Urchin Genes Expressed in Activated Coelomocytes Are Identified by Expressed Sequence Tags

## Complement Homologues and Other Putative Immune Response Genes Suggest Immune System Homology Within the Deuterostomes[1]

L. Courtney Smith,[2,3] Lily Chang, Roy J. Britten, and Eric H. Davidson

**To identify some of the genes expressed in LPS-activated coelomocytes, we sequenced randomly chosen clones from a directionally constructed cDNA library to produce a set of expressed sequence tags (ESTs). Deduced amino acid sequences from 307 ESTs were compared with known protein sequences in GenBank, and significant matches to approximately 30% of the clones were identified. Eighty-nine clones matched to 55 different proteins, including several putative immune effector proteins. In this work, we show the first identification of an invertebrate homologue of a vertebrate C component. Another EST matches to several short consensus repeats that are characteristic of a variety of proteins, including CR/regulatory proteins and clotting factors. Additional putative immune effector genes include 1) a Kazal-type protease inhibitor that may function to inactivate bacterial proteases, 2) a C-type lectin similar to echinoidin, and 3) a serine protease with similarities to thrombin, elastase, haptoglobin, and plasmin. Other EST categories include 1) cell surface proteins and receptors, 2) proteins involved in signaling systems, 3) lysosomal and secreted proteins, 4) cytoskeletal and cytoskeletal modifying proteins, 5) general cell function proteins, 6) proteins with unknown function, and 7) ESTs without significant matches, 25 with open reading frames. Many of the ESTs identified in this study represent the types of genes expected to be used in lower deuterostome immune functions.** ***The Journal of Immunology,*** **1996, 156: 593–602.**

The adaptive immune response of higher vertebrates is a multifaceted, complex, and inter-regulated system. One means to understand it and its evolutionary history is to study its individual facets independent of each other. An approach to accomplish this goal is to investigate organisms with immune systems that function at a simpler level, but yet are phylogenetically related to the higher vertebrates and are classified within the deuterostomes. The echinoderm phylum is the deuterostome sister group of the chordates within which the vertebrates are classified. This position in phylogeny makes the echinoderms an important invertebrate group for comparative studies addressing the evolution of various vertebrate systems, including the immune system. While the immune system of higher vertebrates is composed of many subsystems, immunity in the lower deuterostomes displays a decreased complexity (1). Echinoderm immunity is nonspecific, is based on activation of the immune effector cells, and is mediated by coelomocytes (1–3). The phagocytic coelomocytes show striking similarities with the nonspecific functions of macrophages and leukocytes in higher vertebrates. Adaptive immunity, a hallmark of which is non-self recognition specificity, was added to the underlying nonspecific response of the phagocytes at some point in deuterostome evolution after the divergence of echinoderms. This is based on allograft rejection data showing that sea urchins cannot specifically recognize differences between allografts (1, 4).

The coelomocytes, which are located in the coelomic cavity of the adult sea urchin, mediate the immune response. They consist of several populations, the most prevalent of which is a macrophage-like phagocyte. The activation-type defense functions displayed by the coelomocytes include increases in motility, phagocytosis, encapsulation, and secretion or degranulation of toxic and antibacterial factors (for review, see Ref. 4). In recent studies, we have used increases in profilin transcript level as a marker for coelomocyte activation. Profilin is an actin binding protein that performs a central regulatory role in cytoskeletal reorganization in response to extracellular stimuli (for review, see Ref. 5). By this means, we found that coelomocytes respond to and are activated in vivo by injury (6) and by injections of LPS into the coelom (7). Although sea urchins efficiently clear bacteria from the coelomic cavity (8, 9) and mount an activation-type immune response to both bacteria and bacterial products (7), nothing is yet understood of the genes and molecular mechanisms that are used by these cells in response to immune challenge.

Analysis of expressed sequence tags (ESTs)[4] can quickly yield information on genes expressed in specific cells and tissues (e.g., see Refs. 10–13). In this work, we describe many newly identified genes expressed by activated coelomocytes of the purple sea urchin, *Strongylocentrotus purpuratus*. One of these ESTs represents the first identification of an invertebrate homologue of a vertebrate C gene. Another encodes a C-type lectin with similarities to a

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

Received for publication June 13, 1995. Accepted for publication October 30, 1995.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported by NSF Grant MCB9219330.

[2] Current address: Department of Biological Sciences, George Washington University, Washington, D.C. 20052.

[3] Address correspondence and reprint requests to Dr. L. Courtney Smith, Department of Biological Sciences, George Washington University, Washington, D.C. 20052

[4] Abbreviations used in this paper: EST, expressed sequence tag; SCR, short consensus repeat.



0022-1767/96/$02.00

Table I. *Results of BLAST searches with coelomocyte ESTs*

| Total cDNAs Sequenced | 307 |
|---|---|
| Total matched ESTs | 89 |
| Different genes identified | 55 |
| Other[a] | 56 |
| Unmatched ESTs | 162 |
| Unmatched ESTs with >300 bp | 95 |
| Unmatched ESTs with ORF[b] | 25 |
| Unmatched ESTs without ORF | 70 |

[a] Other includes clones with short inserts or bacterial sequences.
[b] The Grail program (16) was used to identify open reading frames (ORF).

lectin from another sea urchin species and lectin domains identified from cell surface proteins in mammals. Other ESTs show significant similarities to genes encoding proteins involved in putative immune response functions and signaling pathways, in addition to cell surface receptors.

## Materials and Methods

### *cDNA library construction and expressed sequence tag sequence searches*

The sea urchin immune response was activated in vivo by injections of LPS (2 μg LPS/ml coelomic fluid), and coelomocytes from individual animals were collected 24 h later for total RNA isolation and profilin message quantitation by probe excess transcript titration (6, 7). RNA from 18 animals with the highest activation response to LPS was pooled and used to construct a directionally cloned coelomocyte cDNA library. Poly(A) RNA was isolated with oligo(dt) magnetic beads (Dynal, Inc., Great Neck, NY), and cDNA was produced with the Time Saver cDNA kit (Pharmacia, Piscataway, NJ) according to manufacturer's instructions. A random primer containing a *Not*I site (AAA GGA AGG AAA AAA GCG GCC GCT ACA[N]$_8$T) was used to polymerize the first strand, and an *Eco*RI adapter (Pharmacia) was ligated to the 5′ end. After the *Not*I digest, the cDNA was cloned into λ-ExCell phage with *Not*I and *Eco*RI ends (Pharmacia). The library was plated at low density, isolated plaques were picked, and insert sizes were analyzed by PCR using primers to the Sp6 and T7 RNA polymerase sites. Phage with inserts of more than 500 bp were released as phagemids and prepared for sequencing using standard alkaline lysis methods. Cycle sequencing with dye-labeled primers was performed by PCR (Perkin-Elmer Cetus, 9600, Norwalk, CT), and the reactions were analyzed on a 373A automated DNA sequencer (Applied Biosystems, Foster City, CA). EST sequences were translated into all reading frames and used to search the protein sequence data base, Protein Information Resource, with the BLAST program (14). When ESTs were identified containing untranslated sequences, such as 16S rRNA, the FASTA DNA sequence comparison program (15) was used to determine percent identity.

EST match significance resulting from GenBank searches was estimated using an empirically derived method. Test searches of random sequences of 200 amino acids in length, with composition matching to known sea urchin proteins, revealed a distribution of accidental matches of various lengths and percent identity. Very long matches were found with approximately 21% identity, and were a result of the insertions and deletions permitted by the BLASTP search program (14). This background of accidental matching was defined as B. The scoring method depends on the observed reciprocal relationship between length of match (L) and Z, where Z = the observed percent identity, I, minus the background, B. This is expressed as:

$$L \times Z = \text{constant}.$$

In practical form: the quality of fit $Q = L\ (I - B)/\text{constant}$.

The constants in the equation were set so that all of the observed accidental matches had values of $Q \leq 1.0$.

$$Q = L\ (I - 21.2)/739.2.$$

All search results that gave a value of $Q > 1.0$ were examined as being significant matches.

## Results and Discussion

### *Overview of matched expressed sequence tags*

We sequenced the 5′ ends of 307 randomly chosen cDNAs. Of these, 89 showed significant matches to 55 known proteins (Table I) that could be placed into several categories (Table II). These included putative immune effector proteins, cell surface proteins and receptors, proteins involved in signaling systems, secreted proteins, chaperones in the endoplasmic reticulum involved in the secretory machinery, cytoskeletal and cytoskeletal modifying proteins, general cell function proteins, and proteins of unknown function (Table II).

We were interested to determine whether unmatched ESTs, which constituted the largest category (Table I), contained open reading frames or if the sequence for these clones fell in nontranslated regions of the transcript. Using the GRAIL program (16), an analysis of unmatched ESTs of more than 300 bp in length revealed that 25 had open reading frames (Table I). Consequently, this subset of unmatched ESTs represents genes that either have not been previously identified or are composed of sequences too divergent to be identified by sequence comparison programs.

Many ESTs in this study match to genes previously identified in echinoderms, including those in the cytoskeletal, mitochondrial, and general cell function categories. However, others represent newly identified sea urchin genes that fall into the categories of putative immune effector proteins, surface receptors, proteins in signaling pathways, and ribosomal proteins, and match to sequences previously identified in organisms from other phyla.

### *A homologue of complement C3/C4/C5*

The most interesting EST identified from the category of putative immune effector proteins matches to C proteins. This is the first identification of such a homologue in an invertebrate. To verify the initial identification of the short EST064 sequence, additional sequencing was conducted (L. C. Smith and E. H. Davidson, unpublished data). Separate alignment analyses between EST064 and C components C3, C4, C5, and $\alpha_2$-macroglobulin from human and mouse were performed using LFASTA (15) and ClustalV (17), and have been deposited in the EMBL alignment bank (see legend to Table III). The amino acid identities, similarities, and percent identities between EST064 and vertebrate C components that resulted from these alignments are shown in Table III. The LFASTA analysis indicates that the EST064 sequence is equally homologous to C3, C4, C5, and $\alpha_2$-macroglobulin, with amino acid identities ranging from 26.1% to 29.3%. However, alignments using the ClustalV program show fewer amino acid identities and similarities between EST064 and $\alpha_2$-macroglobulin compared with alignments between EST064 and C3, C4, or C5 (Table III). This difference is based on both EST064 and the C components extending beyond the C-terminal end of the $\alpha_2$-macroglobulin sequence. In this region, EST064 shows additional amino acid identities and similarities to the C components. This difference in length and the C-terminal identities can be seen in Figure 1.

The multiple alignment between EST064, C components C3, C4, and C5, and $\alpha_2$-macroglobulin from human and mouse is shown in Figure 1. This alignment covers 492 amino acids of the sea urchin sequence, spans part of the α region of C3, includes most of the γ region of C4, and extends beyond the end of $\alpha_2$-macroglobulin. The portion of the sea urchin protein that has been deduced to date does not include the thioester site that is diagnostic of C3 function; however, it does span the polyarginine αγ junction that is characteristic of C4 proteins. This junction is absent from both EST064 and the other non-C4 proteins (Fig. 1). Inspection of the multiple alignment in Figure 1 reveals that the EST064 sequence matches to the C3 proteins at 124 positions (25.2%), to C4 and C5 at 119 positions (24.1%), and to $\alpha_2$-macroglobulin at 90 positions (18.5%). The EST064 sequence may be directly descended from and similar to the primordial C component theorized to have been present in the ancestral deuterostome, and that would have given rise, by gene duplication, to the C gene family (18).

The identities between the sea urchin sequence and all of the C components used in this analysis indicate how much of the echinoderm sequence has been conserved in the different members of this gene family. When added together, 256 positions of the 492 in the EST064 sequence (52%) match between EST064 and the vertebrate sequences. Taken together, the data presented in Table III and Figure 1 indicate that EST064 is much more distantly related to $\alpha_2$-macroglobulin, and has approximately equal similarity to the vertebrate C components C3, C4, and C5. Future analysis of the cDNA sequence and a biochemical analysis of the protein will identify to which of these three proteins the sea urchin sequence is most closely related.

### *Expressed sequence tag 152 contains short consensus repeats*

BLASTP searches with EST152 revealed matches with proteins that contain multiple short consensus repeats (SCRs), with the most significant matches being to type 1 and type 2 CRs. SCRs are 60 amino acid domains, with common positioning of four cysteines, four glycines, two prolines, a tyrosine (or phenylalanine), and a tryptophane, and they commonly function in protein-protein interactions between C components (19). Examples of C family proteins that match to EST152 include CR1, CR2, factor H, C4 binding protein, C3b/C4b receptor, decay-accelerating factor, C2, and C7. Several non-C proteins containing SCRs were also identified by searches with EST152. These included coagulation factor XIIIb, IL-2R, cell adhesion molecules such as P- and E-selectins, and the lymph node homing receptor (for a review of SCRs, see Ref. 20). The Q value (see Table II) for EST152 is well below significance (Q = 0.27; see *Materials and Methods*), suggesting that the matches between the SCRs in EST152 and other proteins containing SCRs, such as the CR/regulatory proteins, are not real. However, because searches with EST152 repeatedly identified proteins containing SCRs, and because EST152 matches to the consensus sequence for almost four SCRs (Fig. 2), we believe that in this case the low Q value should be ignored. Furthermore, amino acid sequence alignments between the SCRs from other proteins of known function also show identities only at the consensus positions (see Fig. 3 in Ref. 21).

The alignment of the SCRs found in EST152 (after additional sequencing; L. C. Smith and E. H. Davidson, unpublished data) with the consensus sequence is presented in Figure 2 and shows that the *S. purpuratus* sequence spans almost four SCRs. The first and last are partial repeats, and the second matches at all of the consensus positions; however, the third fails to match at the second and third cysteines. The SCR domain is held together by disulfide boding between cysteines 1 and 3, and cysteines 2 and 4. Without all four cysteines, SCR 3 in EST152 will not fold as an SCR. However, the sequence presented in this study for all of the ESTs, including EST152, is imperfect, incomplete, and preliminary. Future work on many of the more interesting ESTs will clarify many questions these preliminary data have raised. In summary, the data presented in this work for EST064 and EST152 suggest that sea urchin coelomocytes express homologues of vertebrate C genes that argue for the presence of a primitive C system that functions in these lower deuterostomes.

### *Possible functions of a primitive echinoderm complement system*

The sea urchin C homologue expressed in coelomocytes shows sequence similarities to vertebrate C proteins (Fig. 1), including the C3-like proteins that have been cloned and sequenced from the cyclostomes (22–24). (A multiple alignment between EST064 and the C3-like proteins from the cyclostomes is available from the EMBL server by E-mailing the request GET ALIGN: DS22986.DAT to Netserv@ebi.ac.uk.) Hagfish C3 functions as a simple opsonin, using the thioester site to bind to yeast that augments the phagocytic activities of hagfish leukocytes to remove the C3-coated yeast (25). An opsonization function similar to that found in the hagfish is predicted for the sea urchin homologue. For an opsonin system to function, not only must there be a molecule that binds to the pathogen (i.e., the opsonin), but there must also be a receptor on the phagocyte that will bind the opsonizing molecule. In vertebrate proteins, SCRs occur in tandem arrays of 2 to 30 repeats, and function in protein-protein interactions between C components, coagulation factors, and cell adhesion molecules, among others (19, 26). Because the sequence of the SCR containing EST from *S. purpuratus* is not only imperfect but also incomplete, and other domains including a transmembrane region have not yet been identified, it is not possible to tell whether EST152 represents a transcript for a CR, a decay-accelerating factor homologue, a cell adhesion protein, or a secreted protein similar to factor B. For comparison, the lamprey C system consists of C3 and factor B, both of which are required for opsonization functions (27, 28). Yet, there is some experimental data to support the presence of a CR in echinoderms. Bertheussen and co-workers (29–32) have shown that phagocytes from *Strongylocentrotus droebachiensis* will take up SRBC coated with mammalian C3b much more readily than red cells without C3b. The authors infer from this data that phagocytes display a receptor for mammalian C3b that augments phagocytosis. Furthermore, the hagfish appears to display a leukocyte surface receptor (33) that may mediate phagocytosis of yeast opsonized by hagfish C3 (24). Thus, the protein encoded by the EST152 gene may be a CR that enhances phagocytosis of opsonized microbial pathogens. Such a simple opsonin system would constitute an important aspect of host defense functions in sea urchins.

An "archeo" C system suggested to have been present in the deuterostome ancestor might have consisted of a single C component, i.e., a primordial C3, and a C binding/regulatory protein (18). This simple system is thought to have given rise, through gene duplication, to the present day C systems that function in higher vertebrates. The newly identified echinoderm system fits the criteria of the "archeo" C system predicted by Lachman (18), and consists of the C3/C4/C5 homologue (EST064) and a protein (EST152) composed, in part, of SCRs. This sea urchin system may be directly descended, with little change, from the simple C system hypothesized to have been present in the deuterostome ancestor.

### *Expressed sequence tags matching other putative immune effector proteins*

***Lectins.*** A variety of defense functions can be inferred from the other ESTs in the category of immune effector proteins. These ESTs match to a lectin, a protease inhibitor, and a serine protease. Lectins are proteins that recognize specific carbohydrates on cell surfaces and extracellular matrices. Two ESTs (EST056/249) match best to C-type ($Ca^{2+}$-dependent) lectins that have 18 invariant residues in a conserved pattern spread throughout a 130-amino-acid region (34). The sequence deduced from EST056 spans 14 of these positions in echinoidin and matches to 12 (alignment not shown). Echinoidin has been characterized from the sea urchin *Anthocidaris crassispina* (35), it recognizes Gal$\beta$1–3GalNAc, and has an RGD sequence that binds to integrins and may function in cell adhesion and spreading (36). Like many other lectins in invertebrates, echinoidin has a single carbohydrate recognition domain, is secreted into the coelomic fluid as dimers and multimers, and may function to opsonize or agglutinate microbes by binding to the surface carbohydrates and thereby augmenting their removal by phagocytic cells (34).

EST056/249 also show sequence similarities to C-type lectin

Table II. *Sea urchin ESTs that match to known proteins*

| EST | Protein Match (Example) | Percentage Identity: Length[a] | Q Value[b] | Genus Match[c] | Kindom, Phylum, or Class[d] | Accession Number |
|---|---|---|---|---|---|---|
| Putative immune effector proteins | | | | | | |
| 064 | Homologue of complement component C3, C4 or C5 | (see Table III) | 9.65 | *Homo* | M | R61937 |
| 152 | Sort consensus repeats (complement receptor or regulatory proteins, factor XIIIb, etc) | 22%:250aa | *0.27 | *Mus* | M | R61992 |
| 056 | Lectin (echinoidin) | 44%:108aa | 3.38 | *Anthocidaris* | E | R61931 |
| | Lectin (brevican) | 35%:109aa | 2.03 | *Bos* | M | |
| 249 | Lectin (echinoidin) | 54%:69aa | 3.06 | *Anthocidaris* | E | R62040 |
| 142 | Serine protease (thrombin, elastase, haptoglobin plasmin) | 40%:52aa | 1.32 | *Bos* | V | R61988 |
| 132 | Protease inhibitor (Kazal type) | 51%:29aa | 1.17 | *Homo* | M | R61984 |
| Cell surface proteins and receptors | | | | | | |
| 059 | Receptor-linked protein tyrosine kinase | 52%:253aa | 10.54 | *Homo* | M | R61934 |
| 369 | High density lipoprotein binding protein | 60%:60aa | 3.15 | *Homo* | M | R62104 |
| 020 | $Na^+K^+$ ATPase, $\alpha$ subunit | 76%:59aa | 4.37 | *Ovis* | M | R61904 |
| 179 | $Na^+K^+$ ATPase, $\alpha$ subunit | 56%:60aa | 2.82 | *Gallus* | V | R62008 |
| 065 | Intestinal brush boarder protein | 40%:62aa | 1.58 | *Oryctolagus* | M | R61938 |
| 166 | Integumentary mucin | 60%:30aa | 1.57 | *Xenopus* | V | R61997 |
| Proteins involved in signaling systems | | | | | | |
| 388 | FKBP-12 | 50%:101aa | 3.94 | *Caenorhabditis* | N | R62119 |
| 317 | GF14 | 76%:55aa | 4.08 | *Arabidopsis* | P | R62074 |
| | 14-3-3 | 41%:55aa | 1.47 | *Caenorhabditis* | N | |
| 030 | Adenylyl cyclase | 61%:36aa | 1.94 | *Rattus* | M | R61912 |
| 112 | Phosphodiesterase ectoenzyme (PC-1, autotoxin) | 69%:26aa | 1.68 | *Rattus* | M | R61966 |
| Lysosomal, secreted or extracellular proteins | | | | | | |
| 003 | Arylsulfatase | 45%:47aa | 1.51 | *Strongylocentrotus* | E | R61895 |
| 004 | Arylsulfatase | 43%:54aa | 1.59 | *Strongylocentrotus* | E | R61896 |
| 072 | Arylsulfatase | 43%:131aa | 3.86 | *Strongylocentrotus* | E | R61942 |
| 401 | Arylsulfatase | 73%:73aa | 5.12 | *Strongylocentrotus* | E | R62129 |
| 118 | Cathepsin S | 40%:57aa | 1.45 | *Oryctolagus* | M | R61972 |
| 052 | Cathepsin L | 87%:16aa | 1.42 | *Homo* | M | R61929 |
| Cytoskeletal and cytoskeletal modifying proteins | | | | | | |
| 196 | $\alpha$-Tubulin | 75%:56aa | 4.08 | *Zea* | P | R62017 |
| 379 | $\alpha$-Tubulin | 80%:62aa | 4.93 | *Drosophila* | Ar | R62111 |
| 284 | $\beta$-Tubulin | 96%:103aa | 10.42 | *Haliotis* | Mo | R62054 |
| 109 | Dynein, heavy chain | 51%:96aa | 3.87 | *Rattus* | M | R61964 |
| 393 | Kinesin, light chain | 93%:32aa | 3.11 | *Strongylocentrotus* | E | R62123 |
| 169 | Actin | 76%:89aa | 6.60 | *Strongylocentrotus* | E | R62000 |
| 194 | Actin | 56%:51aa | 2.40 | *Strongylocentrotus* | E | R62015 |
| 278 | Actin | 94%:106aa | 10.44 | *Strongylocentrotus* | E | R62049 |
| 363 | Actin | 60%:30aa | 1.57 | *Strongylocentrotus* | E | R62100 |
| 400 | Actin | 94%:37aa | 3.64 | *Caenorhabditis* | N | R62128 |
| 009 | Severin, gelsolin | 35%:148aa | 2.76 | *Mus* | M | R61898 |
| 154 | Severin, gelsolin | 39%:53aa | 1.28 | *Dictyostelium* | Y | R61994 |
| 258 | Severin, gelsolin | 43%:55aa | 1.62 | *Dictyostelium* | Y | R62045 |
| 114 | Fascin | 82%:91aa | 7.48 | *Strongylocentrotus* | E | R61968 |
| 074 | Thymosin $\beta_{10}$ | 67%:52aa | 3.22 | *Rattus* | M | R61944 |
| 229 | Rho GDP dissociation inhibitor | 42%:42aa | 1.18 | *Homo* | M | R62034 |
| General cell function proteins | | | | | | |
| 151 | Ferritin, heavy chain | 76%:46aa | 3.41 | *Mus* | M | R61991 |
| 203 | Ferritin, heavy chain | 52%:136aa | 6.50 | *Lymnea* | Mo | R62024 |
| 288 | Ferritin, heavy chain | 78%:14aa | 1.08 | *Mus* | M | R62056 |
| 289 | Ferritin, heavy chain | 75%:60aa | 4.37 | *Homo* | M | R62057 |
| 324 | Ferritin, heavy chain | 54%:46aa | 2.04 | *Rana* | V | R62077 |
| 346 | Ferritin, heavy chain | 54%:97aa | 4.30 | *Lymnea* | Mo | R62088 |
| 153 | Calcium storage protein | 92%:102aa | 9.77 | *Strongylocentrotus* | E | R61993 |
| 042 | Ubiquitin | 100%:20aa | 2.13 | *Oryza* | P | R61923 |
| 321 | Ubiquitin | 97%:77aa | 7.90 | *Drosophila* | Ar | R62076 |
| 395 | Ubiquitin | 71%:35aa | 2.36 | *Strongylocentrotus* | E | R62125 |
| 035 | Mitochondrial protein transporter | 88%:68aa | 6.15 | *Mus* | M | R61917 |
| 398 | Poly(A) binding protein | 61%:91aa | 4.90 | *Mus* | M | R62127 |
| 049 | Heat shock protein (108 kD) | 79%:89aa | 6.96 | *Gallus* | V | R61927 |
| Endoplasmic reticulum chaperones | | | | | | |
| 107 | Sec-61-complex, $\beta$ subunit | 87%:41aa | 3.65 | *Homo* | M | R61962 |
| 102 | Rerlp | 43%:58aa | 1.71 | *Saccharomyces* | Y | R61958 |
| 050 | Heat shock protein-like chaperone (endoplasmin, tumor rejection antigen (gp96) | 75%:24aa | 1.75 | *Mus* | M | R61928 |
| 177 | BiP chaperone | 77%:95aa | 7.17 | *Rattus* | M | R62006 |

(*Continued*)

Table II. *Continued*

| EST | Protein Match (Example) | Percentage Identity: Length[a] | Q Value[b] | Genus Match[c] | Kindom, Phylum, or Class[d] | Accession Number |
|---|---|---|---|---|---|---|
| Ribosomal proteins | | | | | | |
| 061 | Elongation factor 1α | 67%:95aa | 5.89 | *Mucor* | Y | R61936 |
| 054 | Elongation factor 2 | 75%:118 | 8.59 | *Drosophila* | Ar | R61930 |
| 209 | Elongation factor 2 | 40%:124aa | 3.15 | *Caenorhabditis* | N | R62025 |
| 187 | Ribosomal protein L5 | 79%:101aa | 7.90 | *Gallus* | V | R62133 |
| 188 | Ribosomal protein L5 | 80%:60aa | 4.77 | *Gallus* | V | R62013 |
| 265 | Ribosomal protein L6 | 40%:70aa | 1.78 | *Homo* | M | R62046 |
| 135 | Ribosomal protein L7 | 66%:39aa | 2.36 | *Mus* | M | R61986 |
| 219 | Ribosomal protein L8 | 80%:72aa | 5.73 | *Homo* | M | R62029 |
| 104 | Ribosomal protein L10 | 77%:99aa | 7.47 | *Homo* | M | R61959 |
| 171 | Ribosomal protein L10 | 80%:108aa | 8.59 | *Homo* | M | R62002 |
| 309 | Ribosomal protein L32 | 54%:92aa | 4.08 | *Homo* | M | R62069 |
| 325 | Ribosomal protein L37a | 68%:74aa | 4.69 | *Rattus* | M | R62078 |
| 116 | 16S rRNA | 92%:277nt | 8.81 | *Strongylocentrotus* | E | R61970 |
| 170 | 16S rRNA | 92%:420nt | 13.41 | *Strongylocentrotus* | E | R62001 |
| 341 | 16S rRNA | 89%:346nt | 10.55 | *Strongylocentrotus* | E | R62083 |
| Mitochondrial proteins | | | | | | |
| 023 | Cytochrome oxidase 1, subunit 1 | 48%:112aa | 4.06 | *Strongylocentrotus* | E | R67979 |
| 131 | Cytochrome oxidase 1, subunit 1 | 87%:125aa | 11.13 | *Strongylocentrotus* | E | R61983 |
| 235 | Cytochrome oxidase 1, subunit 1 | 91%:93aa | 8.78 | *Strongylocentrotus* | E | R62035 |
| 380 | Cytochrome oxidase 1, subunit 1 | 77%:48aa | 3.62 | *Strongylocentrotus* | E | R62112 |
| 094 | Cytochrome oxidase 1, subunit 3 | 93%:113aa | 10.98 | *Strongylocentrotus* | E | R61955 |
| 189 | Cytochrome oxidase 1, subunit 3 | 90%:55aa | 5.12 | *Strongylocentrotus* | E | R62914 |
| 285 | Cytochrome b | 76%:59aa | 4.37 | *Strongylocentrotus* | E | R62055 |
| 308 | Cytochrome b | 96%:46aa | 4.65 | *Strongylocentrotus* | E | R62068 |
| 397 | Cytochrome b | 72%:22aa | 1.51 | *Strongylocentrotus* | E | R62126 |
| 225 | Cytochrome C oxidase, subunit VIb | 68%:38aa | 2.41 | *Bos* | V | R62031 |
| 226 | Cytochrome C oxidase, subunit VIb | 66%:56aa | 3.39 | *Bos* | V | R62032 |
| 036 | ATP synthase, α-subunit | 96%:110aa | 11.13 | *Xenopus* | V | R61918 |
| 093 | ATPase 6 | 60%:56aa | 2.94 | *Strongylocentrotus* | E | R61954 |
| 085 | NADH dehydrogenase, subunit 5 | 83%:31aa | 2.59 | *Strongylocentrotus* | E | R61952 |
| 361 | NADH dehydrogenase, subunit 5 | 88%:36aa | 3.25 | *Strongylocentrotus* | E | R62098 |
| Proteins of unknown function | | | | | | |
| 186 | Csa-19 | 59%:87aa | 4.45 | *Homo* | M | R62012 |
| 178 | Arabidopsis EST Z26403 | 62%:61aa | 3.37 | *Arabidopsis* | P | R62007 |
| 183 | Arabidopsis EST Z26403 | 63%:52aa | 2.94 | *Arabidopsis* | P | R62010 |
| 157 | Maternal nontranslated message | 84%:108aa | 9.18 | *Strongylocentrotus* | E | R61995 |

[a] The percentage identity of amino acids (aa) between the ESTs and the protein to which they matched best were calculated with the BLAST program (14).
[b] The Q value, or match significance, was calculated according to the equation given in the methods section for each EST. Those ESTs with $Q \geq 1.0$ were considered significant with the exception of EST152. For those ESTs that matched to 16S rRNA, the percentage identity was determined with the LFASTA program (15) and the Q value was calculated by a variation of the equation given in the *Materials and Methods* section based on the percentage identity and 1/3 the length of nucleotides (nt): $Q = (L/3)(I - 21.2)/739.2$. *, the Q value for EST152 is below a significant value. See text for discussion.
[c] Th best matches are listed by genus.
[d] The kingdom, phylum, or class is listed by abbreviation: M, mammal; V, vertebrate; E, echinoderm; Ar, arthropod; Mo, mollusc; An, annelid; N, nematode; Y, yeast or fungus; P, plant.

Table III. *Amino acid sequence comparisons between EST064 and several vertebrate complement components*

| EST064 Alignment with: | LFASTA | Clustal V[a] | | |
|---|---|---|---|---|
| | % Identity:Length of amino acids | Amino acid identities to both vertebrate sequences | Amino acid similarities (and matches to one vertebrate protein) | Number of amino acids from the vertebrate sequences included in the alignments |
| C3, human | 27.6%:377 | 111 | 118 (18) | 521 |
| C3, mouse | 26.2%:393 | | | |
| C4, human | 28.4%:208 | 117 | 141 (12) | 579 (includes αγ hinge region) |
| C4, mouse | 29.3%:140 | | | |
| C5, human | 26.1%:464 | 119 | 129 (14) | 513 |
| C5, mouse | 26.8%:455 | | | |
| a2m*, human | 28.0%:347 | 86 | 82 (22) | 380 |
| a2m, mouse | 27.6%:341 | | | |

[a] Multiple alignments were done between EST064 and both human and mouse sequences using ClustalV and the identities and similarities were counted. The alignments are available from the EMBL server by e-mail. Send requests to Netserv@ebi.ac.uk and include GET ALIGN:DSXXXXX.DAT in the message. The data submission number for each alignment (to be inserted in place of the Xs) are as follows: EST064 aligned with C3, 22983; with C4, 22984; with C5, 22985; with α-2 macroglobulin, 22987. *, the number of identities and similarities between EST064 and α-2 macroglobulin (a2m) is substantially lower than between EST064 and the other complement components. This is because a2m is shorter than the other proteins and alignments between EST064 and the other complement components extend beyond the C-terminal end of a2m.

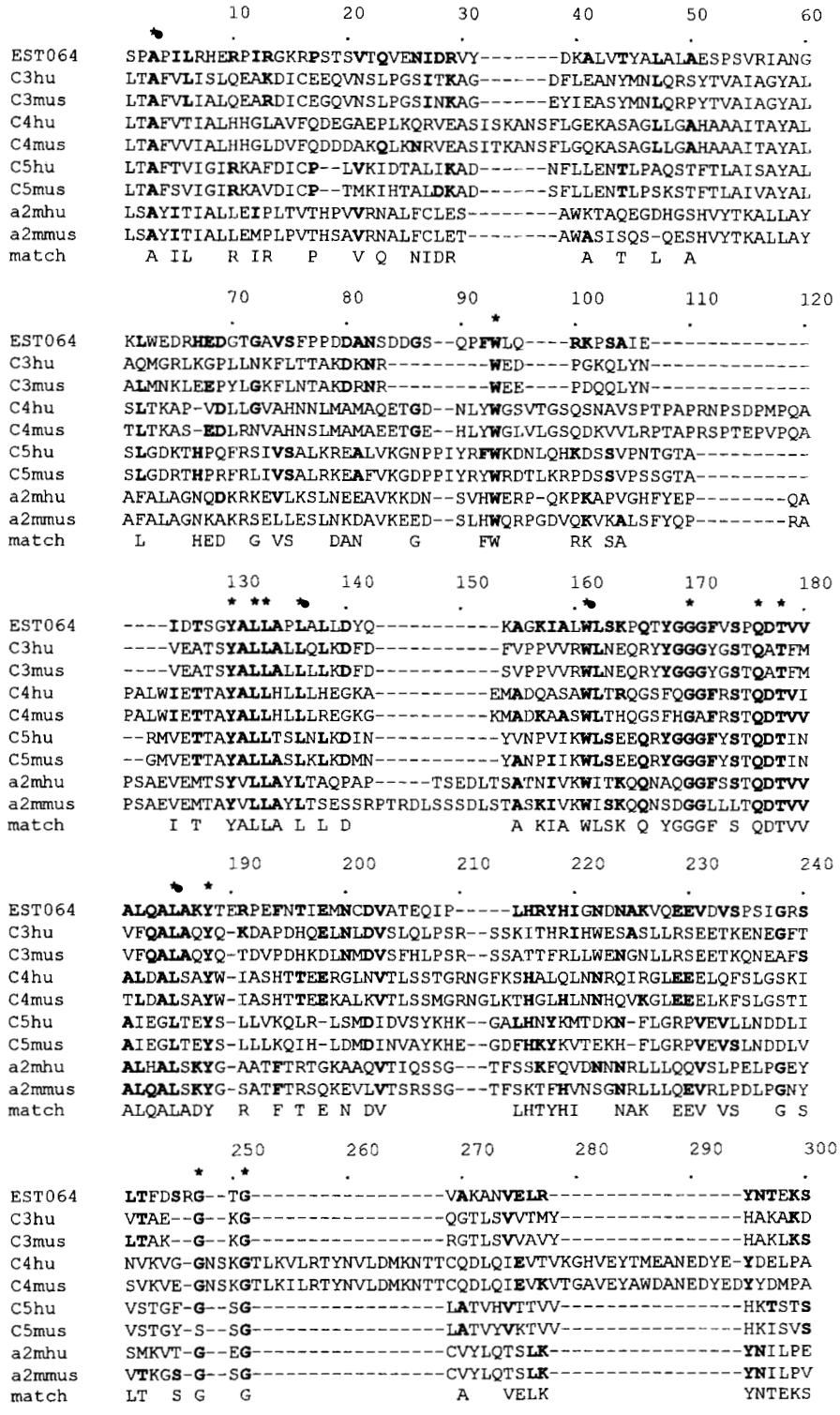


**FIGURE 1.** Alignment of coelomocyte EST064 with human and mouse C components. This alignment was performed using ClustalV (17) without subsequent adjustments by hand. Because of the incomplete sequence for EST064, the alignment covers the carboxyl-terminal portions of the mammalian proteins, beginning at approximately the middle of the C3 α-chain. All positions that match between EST064 and any of the mammalian sequences are denoted in bold and are listed as a match on the bottom line. Lysine and arginine are considered to be a match. Positions that match in all sequences are identified with a star above the position. The stretch of arginines (italicized) in the C4 sequences indicates the αγ junction. hu, human; mus, mouse; a2m, $\alpha_2$-macroglobulin.

domains found in multidomain proteins from mammals that include (in order of match significance) brain proteins of the neurocan/brevican family (37), the low affinity IgE-Fc receptor or CD23 (38), and the macrophage mannan receptor (39). Lectin domains in multidomain proteins can have a variety of functions. For example, the neurocan/brevican protein family consists of large, multifunctional extracellular brain-specific proteins that contain a single C-type lectin domain that may function in interactions between cell surface carbohydrates and hyaluronic acid (37). The single lectin domain in CD23 retains its lectin structure, but it has lost its carbohydrate-binding capabilities, and instead binds the Fc portion of IgE (38). The macrophage mannan receptor, on the other hand, has eight extracellular C-type lectin domains, is constitutively internalized and recycled (38), and mediates phagocytosis of mannose-coated cells and glycoproteins, including the uptake of bacterial pathogens (40). This is the first lectin identified in *S. purpuratus*, and whether it is a secreted lectin or a lectin domain within a larger protein, it may function as an opsonin system, in addition to

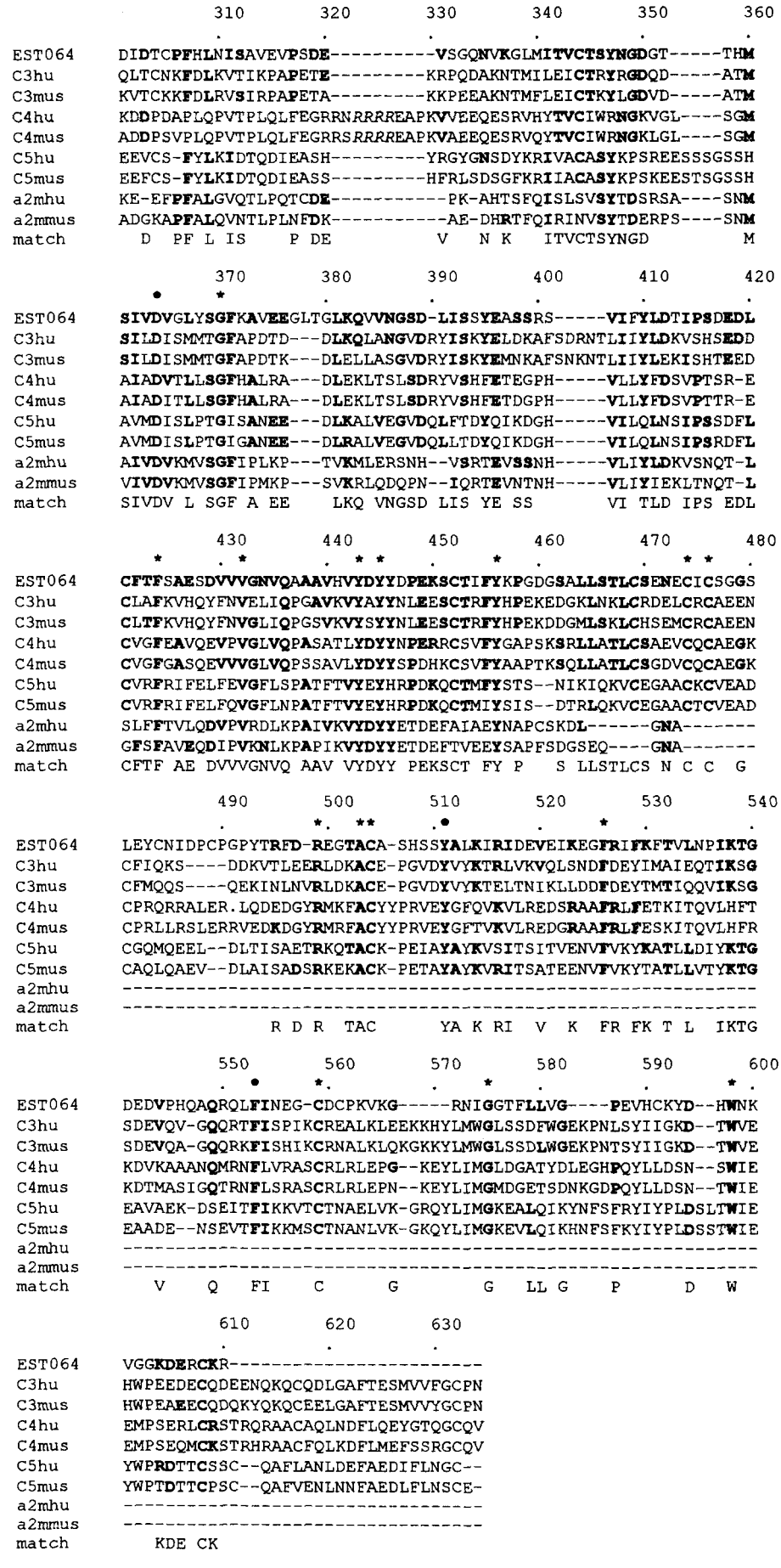


**FIGURE 1.** *Continued*

```
SCR1          RHEGKTLGSNTVISHTERTVFNYTCQT-GYHPTGDVTITCQ-----NGS-------WEADADVPACNKII
SCR2      CAPPTLEPIENGTHNIENRTYNVGEMVQYECNE-GYLVDGISWQVCQSHDGPNGI-------WSDNV--PRCKVAR
SCR3      C-PHSGFPSPQRRDFFLAEKRTRRHE--FLARRSGDDPQVSVTSRAAWVKE--GVSTRWSLEWYSDVM-PQCIEYD
SCR4      C-PALGDPXHGSXVGSPQXXVNTKVH--FRCIH-GCPMVGSDSXGCVRFLXXW
consensus C PP                      Y/F C  G  G  C  G  W  P C
```

**FIGURE 2.** Alignment of the SCR from coelomocyte EST152. This alignment was performed initially with the ClustalV (17) program, but had substantial adjustments by hand. The EST152 sequence spans almost four SCRs; the first and last repeats are partial due to the incomplete sequence. The consensus positions in SCRs from Dierich et al. (19) are shown on the bottom line. The fourth consensus position can be either tyrosine or phenylalanine. X denotes an unknown amino acid in the EST152 sequence resulting from an ambiguous nucleotide.

*S. purpuratus* C homologue, for microbial agglutination and phagocytosis.

***Protease inhibitors.*** A common invertebrate defense mechanism involves protease inhibitors. These proteins are secreted into the coelomic cavity in response to infection, and function to inactivate microbial proteases and thereby counteract the spread of infections (41). EST132 matches to two of the several Kazal-type protease inhibitor domains that are part of a large, multidomain protein found in neuromuscular junctions called agrin (42; for review, see Ref. 43). Serine protease inhibitors fall into 10 or more families, the Kazal type being one of the better studied, and are based on the topologic relationship between the disulfide bridge arrangement and the reactive site (44). Although we do not expect the protein encoded by the EST132 gene to be a homologue of agrin, it is likely to be a typical protease inhibitor similar in function to other protease inhibitors that have been identified in a sea cucumber (45), and several other invertebrates.

***Proteases.*** The sequence obtained from EST142 matches to several serine proteases that include, in order of match significance, thrombin, elastase, haptoglobin, and plasmin. Thrombin and plasmin are members of the coagulation and fibrinolytic cascades, and have opposite activities on fibrin (46). Thrombin initiates fibrin clot formation by cleaving fibrinogen to form fibrin that is then cross-linked into clots, while plasmin, on the other hand, has fibrinolytic activity and dissolves clots (47). Elastase degrades extracellular matrices (48) and is also involved in the clotting cascade (49) by cleaving and activating plasminogen (50). Haptoglobin is composed of a serine protease domain plus two SCRs (51), and is an acute phase response protein (52). It is secreted by hepatocytes along with other acute phase proteins, including C3, after stimulation by inflammatory cytokines (53). The common denominator in this set of serine proteases is that they are involved in the clotting cascade, the fibrinolytic cascade, and the acute phase response, all of which are aspects of mammalian inflammation. Furthermore, thrombin exerts a variety of activating functions on cells at sites of vascular injury (54, 55). Elastase, secreted from monocytes and neutrophils at sites of inflammation, has antimicrobial activity that aids in the local degradation of foreign microbes (56). Serine proteases characterized from invertebrates also function in inflammatory responses to injury and infection, and include the prophenyloxidase cascade in tunicates (57) and the clotting cascade of *Limulus* (58). This common theme of inflammatory responses involving serine proteases suggests that the *S. purpuratus* serine protease encoded by the EST142 gene may have a similar function. It may be secreted into the coelomic fluid by the coelomocytes and be involved in the sea urchin inflammatory responses, perhaps by regulating clotting functions or by degrading microbial pathogens.

### *Expressed sequence tags encoding other proteins: implications for other coelomocyte functions*

Additional functions and activities of coelomocytes can be inferred from other ESTs listed in Table II. Several ESTs show similarities to cell surface receptors and proteins involved in signaling pathways. EST059 matches to a growth factor receptor containing a protein tyrosine kinase domain in the cytoplasmic tail. Best matches to EST059 include the fibroblast growth factor receptor and the platelet-derived growth factor receptor. EST388 shows sequence similarity to FKBP-12 and FKBP-13. FKBP-12 has been identified in mammalian cells and yeast (59), and is part of a signaling pathway that includes a cell surface receptor that has similar sequence in both mammals and yeast (60, 61), a cytoplasmic phosphatase called calcineurin (62–64), and several ubiquitous transactivating DNA binding proteins, including nuclear factor-$\kappa$B (65). In T cells, activation of the FKBP12 pathway results in cell activation (62) and cell proliferation (60), as mediated by the DNA binding protein nuclear factor-ATP (66).

EST317 shows sequence similarities to a set of proteins that have been identified in vertebrates, insects, yeast, and plants, called 14–3–3 or GF14 (67). These cytoplasmic proteins are involved in a different signaling pathway from that discussed above for EST388, and function by activating tyrosine and tryptophane hydroxylases (68) and regulating kinase-mediated signal transduction pathways (69, 70). In addition, 14–3–3 may also be involved in host responses to pathogens (67), perhaps by $Ca^{2+}$-mediated exocytosis (71). GF14, in plants, may be part of the DNA binding complex that interacts at G box sequences (70, 72), providing regulatory signals important in environmental responsiveness (72). The expression of EST059, EST388, and EST317 in coelomocyte suggests that coelomocytes may have a variety of methods for sampling, responding to, and maintaining the homeostasis of the coelomic environment.

EST030 and EST112 have sequence similarities to adenylyl cyclase and a phosphodiesterase domain, respectively. Cyclases are involved in modulating cyclic nucleotide concentrations that affect second messenger systems. EST112 is similar to the phosphodiesterase domain of the plasma cell differentiation Ag-1 (PC-1) and autotaxin. The phosphodiesterase domains of both PC-1 and autotaxin are located in the extracellular region of these cell surface proteins (73, 74), and both function as ectoenzymes. Their cell surface activities include stimulating cell motility, interactions between cells and the extracellular matrix (74), and may have cell signaling capabilities through a G protein pathway (73). The putative expression of ectoenzymes on coelomocyte surfaces suggests the presence of another system, besides the more classic cell surface receptors discussed above, through which these cells may receive and respond to extracellular information.

Coelomocytes are activated easily by injury and infection, and respond with increased rates of motility, secretion, and phagocytosis (4, 75), all of which require modifications to the cytoskeleton. These activities are reflected by the number of ESTs that match to proteins involved in the cytoskeleton, and include actin, gelsolin, fascin, and thymosin, among others. Expression of $\alpha$ and $\beta$ tubulin (EST196/379 and EST284) and dynein (EST109) would be expected in the vibratile cells, which have a single flagellum. Active secretory processes are suggested by the identification of protein chaperones found in the endoplasmic reticulum that are involved in directing proteins through the secretory pathway (EST107 and

EST102). In addition, several secreted proteins were identified. Arylsulfatase (EST003/004/072/401) is localized in lysosomes of mammalian macrophages and is expressed by the aboral ectoderm of sea urchin embryos (76). In mammals, arylsulfatase functions extracellularly to hydrolyze sulfated glycolipids, glycoproteins, and oligosaccharides (77). In sea urchin embryos, arylsulfatase is secreted into the blastocoel and may function to interconnect sulfated components of the extracellular matrix (78). Arylsulfatase has also been identified biochemically in sea urchin coelomocytes (79) and may be released from granules by exocytosis during cytotoxic reactions (80). Cathepsin L (EST052) and cathepsin S (EST118) are cysteine proteases found in macrophage lysosomes in mammals, and function to degrade basement membranes and the extracellular matrix (81, 82). Secretion of arylsulfatase, the cathepsins, and perhaps the serine protease discussed above suggests that coelomocytes have the capabilities to invade tissues from the coelomic fluid, thus allowing them to perform surveillance or patrolling functions to identify and remove points of infection, or to remedy tissue damage.

### *Newly identified genes expressed in coelomocytes: implications for the evolution of deuterostome immunity*

We have identified many new genes expressed by the *S. purpuratus* coelomocytes, and thus, we are able to predict some of the functions and activities that these cells perform in the adult sea urchin. This study also illustrates several points regarding the evolution of the deuterostome immune response. A central question regarding the evolution of immunity in higher vertebrates has been one of origins: when and how did the adaptive immune response originate? What were the first steps in the evolutionary process that resulted in adaptive immunity? Was the occurrence of a simple, component-based opsonin system involved? The evolutionary expansion of the primordial components by gene duplication (18, 26, 83) into a cytolytic system appears to have occurred concurrently with the invention of the Ig system in the elasmobranchs (see Ref. 4, Fig. 1). The cyclostomes do not appear to express Ig (searches for Ig have gone on since the 1970s without success) and have been shown, to date, to have a limited component system (21, 23, 28). However, the presence of a primitive compliment or C system composed of one component and a receptor functioning as an opsonin system may have been the required platform onto which the Ig gene functions were added that resulted in the expansion of both systems in the higher vertebrates. Continued analysis of many of the newly identified genes in *S. purpuratus* will further our understanding of the mechanisms by which the sea urchin immune response works, and will also help define the primitive functions that still operate in the higher vertebrates.

## Acknowledgments

The authors thank Drs. E. Rothenberg and M. Kroninberg and the two anonymous reviewers for their valuable suggestions.

## References

1. Smith, L. C., and E. H. Davidson. 1992. The echinoid immune system and the phylogenetic occurrence of immune mechanisms in deuterostomes. *Immunol. Today 13:356.*
2. Coffaro, K. A., and R. T. Hinegardner. 1977. Immune response in the sea urchin *Lytechinus pictus. Science 197:1389.*
3. Coffaro, K. A. 1979. Transplantation immunity in the sea urchin. Doctoral dissertation, University of California, Santa Cruz, CA.
4. Smith, L. C., and E. H. Davidson. 1994. The echinoderm immune system: characters shared with vertebrate immune systems, and characters arising later in deuterostome phylogeny. In *Primordial Immunity: Foundations for the Vertebrate Immune System.* G. Beck, E. L. Cooper, G. S. Habicht, and J. J. Marchalonis, eds. New York Academy of Sciences, New York, *712:213–226.*
5. Sohn, R. H., and P. J. Goldschmidt-Clermont. 1994. Profilin: at the crossroads of signal transduction and the actin cytoskeleton. *Bioessays 16:465.*
6. Smith, L. C., R. J. Britten, and E. H. Davidson. 1992. *SpCoel1*, a sea urchin profilin gene expressed specifically in coelomocytes in response to injury. *Mol. Biol. Cell 3:403.*
7. Smith, L. C., R. J. Britten, and E. H. Davidson. 1995. Lipopolysaccharide activates the sea urchin immune system. *Dev. Comp. Immunol. 19:217.*
8. Yui, M. A., and C. J. Bayne. 1983. Echinoderm immunology: bacterial clearance by the sea urchin *Strongylocentrotus purpuratus. Biol. Bull. 165:473.*
9. Plytycz, B., and R. Seljelid. 1993. Bacterial clearance by the sea urchin, *Strongylocentrotus droebachiensis. Dev. Comp. Immunol. 17:283.*
10. Adams, M. D., A. R. Kerlavage, C. Fields, and J. C. Venter. 1993. 3,400 new expressed sequence tags identify diversity of transcripts in human brain. *Nat. Genet. 4:256.*
11. Hofte, H., T. Desprez, J. Amselem, H. Chiapello, P. Rouze, and M. Caboche. 1993. An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana. Plant J. 4:1051.*
12. Umeda, M., C. Hara, Y. Matsubayashi, H. H. Li, Q. Liu, F. Tadokoro, S. Aotsuka, and H. Uchimiya. 1994. Expressed sequence tags from cultured cells of rice (*Oryza sativa* L.) under stressed conditions: analysis of transcripts of genes engaged in ATP-generating pathways. *Plant Mol. Biol. 25:469.*
13. Brettini, E., A. R. Prota, N. Dahmen, H. Wang, and F. L. Margolis. 1994. Expressed sequence tags (EST) identify genes preferentially expressed in catfish chemosensory tissues. *Mol. Brain Res. 23:285.*
14. Altschul, S., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol. 215:403.*
15. Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA 85:2444.*
16. Uberbacher, E., and R. Mural. 1991. Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. *Proc. Natl. Acad. Sci. USA 88:11261.*
17. Higgins, D. G., A. J. Bleasby, and R. Fuchs. 1992. Clustal V: improved software for multiple sequence alignment. *Comput. Appl. Biosci. 8:189.*
18. Lachman, P. J. 1979. An evolutionary view of the complement system. *Behring Inst. Mitt. 63:25.*
19. Dierich, M. P., T. F. Schultz, A. Eigentler, H. Huemer, and W. Schwäble. 1988. Structural and functional relationships among receptors and regulators of the complement system. *Mol. Immunol. 25:1043.*
20. Reid, K. B. M., and A. J. Day. 1989. Structure-function relationships of the complement components. *Immunol. Today 10:177.*
21. Zipfel, P. F., and C. Skerka. 1994. Complement factor H and related proteins: an expanding family of complement-regulatory proteins? *Immunol. Today 15:121.*
22. Hanley, P. J., J. W. Hook, D. A. Raftos, A. A. Gooley, R. Trent, and R. L. Raison. 1992. Hagfish humoral defense protein exhibits structural and functional homology with mammalian complement components. *Proc. Natl. Acad. Sci. USA 89: 7910.*
23. Ishiguro, H., K. Kobayashi, M. Suzuki, K. Titani, S. Tomonaga, and Y. Kurosawa. 1992. Isolation of a hagfish gene that encodes a complement component. *EMBO J. 11:829.*
24. Nonaka, M., and M. Takahashi. 1992. Complete complementary DNA sequence of the third component of complement of lamprey. *J. Immunol. 148:3290.*
25. Raftos, D. A., J. W. Hook, and R. L. Raison. 1992. Complement-like protein from the phylogenetically primitive vertebrate, *Eptatretus stouti*, is a homoral opsonin. *Comp. Biochem. Physiol. 103:379.*
26. Farries, T. C., and J. P. Atkinson. 1991. Evolution of the complement system. *Immunol. Today 12:295.*
27. Nonaka, M., R. Fugii, T. Kaidoh, S. Natsuume-Sadai, N. Yamaguchi, and M. Takahashi. 1984. Purification of a lamprey complement protein homologous to the third complement of the mammalian complement system. *J. Immunol. 133:3242.*
28. Nonaka, M., M. Takahashi, and M. Sasaki. 1994. Molecular cloning of a lamprey homologue of the mammalian MHC class III gene, complement factor B. *J. Immunol. 152:2263.*
29. Bertheussen, K. 1981. Endocytosis by echinoid phagocytes in vitro. I. Recognition of foreign matter. *Dev. Comp. Immunol. 5:241.*
30. Bertheussen, K. 1982. Receptors for complement on echinoid phagocytes. II. Purified human complement mediates echinoid phagocytosis. *Dev. Comp. Immunol. 6:635.*
31. Bertheussen, K. 1983. Complement-like activity in sea urchin coelomic fluid. *Dev. Comp. Immunol. 7:21.*
32. Bertheussen, K., and R. Seljelid. 1982. Receptors for complement on echinoid phagocytes. I. The opsonic effect of vertebrate sera on echinoid phagocytosis. *Dev. Comp. Immunol. 6:42 3.*
33. Raison, R. L., J. Coverley, J. W. Hood, P. Towns, K. M. Weston, and D. A. Raftos. 1994. A cell-surface opsonic receptor on leukocytes from the phylogenetically primitive vertebrate, *Eptatretus stouti. Immunol. Cell Biol. 72:326.*
34. Drickamer, K. 1988. Two distinct classes of carbohydrate-recognition domains in animal lectins. *J. Biol. Chem. 263:9557.*
35. Giga, Y., A. Ikai, and K. Takahashi. 1987. The complete amino acid sequence of echinoidin, a lectin from the coelomic fluid of the sea urchin *Anthocidaris crassispina. J. Biol. Chem. 262:6197.*
36. Ozeki, Y., T. Matsui, and K. Titani. 1991. Cell adhesive activity of two animal lectins through different recognition mechanisms. *FEBS Lett. 289:145.*
37. Yamada, H., K. Watanabe, M. Shimonaka, and Y. Yamaguchi. 1994. Molecular cloning of brevican, a novel brain proteoglycan of the aggrecan/versican family. *J. Biol. Chem. 269:10119.*
38. Hajela, K. 1990. Were lectins primitive Fc receptors? *Immunol. Lett. 27:183.*
39. Stahl, P. D. 1992. The mannose receptor and other macrophage lectins. *Curr. Opin. Immunol. 4:49.*

40. Ezekowitz, R. A. F., D. J. Williams, H. Koziel, M. Y. K. Armstrong, A. Warner, F. F. Richards, and R. J. Rose. 1991. Uptake of *Pneumocystis carinii* mediated by the macrophage mannose receptor. *Nature 351:155.*
41. Eguchi, M. 1993. Protein protease inhibitors in insects and comparison with mammalian inhibitors. *Comp. Biochem. Physiol. 105B:449.*
42. Rupp, F., D. G. Payan, C. Magill-Solc, D. M. Cowan, and R. H. Scheller. 1991. Structure and expression of a rat agrin. *Neuron 6:811.*
43. Patthy, L., and K. Nikolics. 1993. Functions of agrin and agrin-related proteins. *Trends Neurosci. 16:76.*
44. Kumazaki, T., and S.-I. Ishii. 1990. Disulfide bridge structure of ascidian trypsin inhibitor I: similarity to Kazal-type inhibitors. *J. Biochem. 107:414.*
45. Matsui, T., Y. Ozeki, M. Suxuki, A. Hino, and K. Titani. 1994. Purification and characterization of two $Ca^{2+}$-dependent lectins from celome plasma of sea cucumber, *Stichopus japonicus. J. Biochem. 116:1127.*
46. Shih, G. C., and K. A. Hajjar. 1993. Plasminogen and plasminogen activator assembly on the human endothelial cell. *Proc. Soc. Exp. Biol. Med. 202:258.*
47. Stubbs, M. T., and W. Bode. 1994. Coagulation factors and their inhibitors. *Curr. Opinion Struct. Biol. 4:823.*
48. Dortho, M. P., P. H. Jarreau, C. Delacourt, I. Macquinmavier, M. Levame, S. Pezet, A. Harf, and C. Lafuma. 1994. Matrix metalloproteinase and elastase activities in LPS-induced acute lung injury in guinea pigs. *Am. J. Physiol. 266: L209.*
49. Anderssen, T., H. Halvorsen, S. P. Bajaj, and B. Osterud. 1993. Human leukocyte elastase and cathepsin-G inactivate factor-VII by limited proteolysis. *Thromb. Haemost. 70:414.*
50. Nakajima, K., K. Nagata, M. Hamanoue, N. Takemoto, and S. Kohsaka. 1993. Microglia-derived elastase produces a low-molecular-weight plasminogen that enhances neurite ourgrowth in rat neocortical explant cultures. *J. Neurochem. 61:2155.*
51. Nicholsonweller, A., J. Zaia, M. G. Raum, and J. E. Colegan. 1987. Decay acceleration factor (DAF) peptide sequence shares homology with a consensus sequence found in the superfamily of structurally related complement proteins and other proteins including haptoglobin, factor XIII, $\beta$2 glycoprotein I, and the IL-2 receptor. *Immunol. Lett. 14:307.*
52. Mackiewicz, A., H. Schooltink, P. C. Heinrich, and S. Rose-John. 1992. Complex of soluble human IL-6-receptor/IL-6 up-regulates expression of acute-phase proteins. *J. Immunol. 149:2021.*
53. Baumann, H., K. K. Morella, and G. H. W. Wong. 1993. TNF-a, IL-1$\beta$, and hepatocyte growth factor cooperate in stimulating specific acute phase plasma protein genes in rat hepatoma cells. *J. Immunol. 151:4248.*
54. Shuman, M. A. 1986. Thrombin-cellular interactions. *Ann. NY Acad. Sci. 485: 228.*
55. Grandaliano, G., A. J. Valente, and H. E. Abboud. 1994. A novel biologic activity of thrombin: stimulation of monocyte chemotactic protein production. *J. Exp. Med. 179:1737.*
56. Gabay, J. E. 1994. Antimicrobial proteins with homology to serine proteases. *Ciba Found. Symp. 186:237.*
57. Jackson, A. D., V. J. Smith, and C. M. Peddie. 1993. In vitro phenoloxidase activity in the blood of *Ciona intestinalis* and other ascidians. *Dev. Comp. Immunol. 17:97.*
58. Muta, T., F. Tokunaga, and T. Nakamura. 1993. *Limulus* clotting factor-C: lipopolysaccharide-sensitive serine protease zymogen. *Methods Enzymol. 223:336.*
59. Brown, E. J., M. W. Albers, T. B. Shin, K. Ichikawa, C. T. Keith, W. S. Lane, and L. S. Schreiber. 1994. A mammalian protein targeted by $G_1$-arresting rapamycin-receptor complex. *Nature 369:756.*
60. Eng, W. K., L. Faucette, and M. M. McLaughlin. 1994. The yeast FKS1 gene encodes a novel membrane-protein, mutations in which confer FK506 and cyclosporin A hypersensitivity and calcineurin-dependent growth. *Gene 151:61.*
61. Sabers, C. J., M. M. Martin, and G. J. Brunn. 1995. Isolation of a protein target of the FKBP12-rapamycin complex in mammalian cells. *J. Biol. Chem. 270:815.*
62. Breuder, T., C. S. Hemenway, N. R. Movva, M. E. Cardenas, and J. Heitman. 1994. Calcineurin is essential in cyclosporin-A-sensitive and FK506-sensitive yeast strains. *Proc. Natl. Acad. Sci. USA 91:5372.*
63. Sabatini, D. M., H. Erdjumentvromage, M. Lui, P. Tempst, and S. H. Snyder. 1994. RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORS *Cell 78:35.*
64. Fruman, D. A., B. E. Bierer, J. Benes, S. J. Burakoff, K. F. Austen, and H. R. Katz. 1995. The complex of FK506-binding protein 12 and FK506 inhibits calcineurin phosphatase activity and IgE activation-induced cytokine transcripts, but not exocytosis, in mouse mast cells. *J. Immunol. 154:1846.*
65. Venkataraman, L., S. J. Burakoff, and R. Sen. 1995. FK506 inhibits antigen receptor-mediated induction of c-*rel* in B and T lymphoid cells. *J. Exp. Med. 181:1091.*
66. Rao, A. 1994. NF-ATP: a transcription factor required for the coordinate induction of several cytokine genes. *Immunol. Today 15:274.*
67. Aitken, A., D. B. Collinge, B. P. H. van Heusden, T. Isobe, P. H. Roseboom, G. Rosenfeld, and J. Soll. 1992. 14–3–3 proteins: a highly conserved, widespread family of eukaryotic proteins. *TIBS17:498.*
68. Ichimura, T., T. Isobe, T. Okuyama, T. Yamauchi, and H. Fujisawa. 1987. Brain 14–3–3 protein is an activator protein that activates tryptophan 5-monooxygenase and tyrosine 3-monooxygenase in the presence of a $Ca^{2+}$ calmodulin-dependent protein kinase II. *FEBS Lett. 219:79.*
69. Hirsch, S., A. Aitken, U. Bertsch, and J. Soll. 1992. A plant homologue to mammalian brain 14–3–3 protein and protein kinase C inhibitor. *FEBS Lett. 296:222.*
70. De Vetten, N. C., G. Lu, and R. J. Ferl. 1992. A maize protein associated with the G-box binding complex has homology to brain regulatory proteins. *Plant Cell 4:1295.*
71. Brandt, J., H. Thordal-Christensen, K. Vad, P. L. Gregersen, and D. B. Collinge. 1992. A pathogen-induced gene of barley encodes a protein showing high similarity to a protein kinase regulator. *Plant J. 2:815.*
72. Lu, G., P. C. Sehnke, and R. J. Ferl. 1994. Phosphorylation and calcium binding properties of an *Arabidopsis* GF14 brain protein homolog. *Plant Cell 6:501.*
73. Buckley, M. F., K. A. Loveland, W. J. McKinstry, O. M. Garson, and J. W. Goding. 1990. Plasma cell membrane glycoprotein PC-1. *J. Biol. Chem. 265: 17506.*
74. Murata, J., H. Y. Lee, T. Clair, H. C. Krutzsch, A. A. Arestak, M. E. Sobel, L. A. Liotta, and M. L. Stracke. 1990. cDNA cloning of the human tumor motility-stimulating protein, autotaxin, reveals a homology with phosphodiesterases. *J. Biol. Chem. 269:30479.*
75. Johnson, P. T. 1969. The coelomic elements of sea urchins (*Strongylocentrotus*). I. The normal coelomocytes; their morphology and dynamics in hanging drops. *J. Invertebr. Pathol. 13:25.*
76. Yang, Q., L. M. Angerer, and R. C. Angerer. 1989. Structure and tissue-specific developmental expression of a sea urchin arylsulfatase gene. *Dev. Biol. 135:53.*
77. Nicholls, R. G., and A. B. Roy. 1971. Arylsulfatases. In *The Enzymes*, Vol 5. P. D. Boyer, ed. Academic Press, New York, p. 21.
78. Sasaki, H., K. Akasaka, H. Shimada, and T. Shimada. 1987. Purification and characterization of arylsulfatase from sea urchin (*Hemicentrotus pulcherrimus*) embryos. *Comp. Biochem. Physiol. 88B:147.*
79. Canicatti, C., and A. Miglietta. 1989. Arylsulfatase in echinoderm immunocompetent cells. *Histochem. J. 21:419.*
80. Pagliara, P., and C. Canicatti. 1993. Isolation of cytolytic granules from sea urchin amoebocytes. *Eur. J. Cell Biol. 60:179.*
81. Delaisse, J. M., P. Ledent, and G. Vaes. 1991. Collagenolytic cysteine proteinases of bone tissue: cathepsin-B, (pro)cathepsin-L and cathepsin-L-like 70-kDa proteinase. *Biochem. J. 279d:167.*
82. Shi, G. P., J. S. Munger, J. P. Meara, and H. A. Chapman. 1992. Molecular cloning and expression of human alveolar macrophage cathepsin-S, an elastinolytic cysteine protease. *J. Biol. Chem. 276:7258.*
83. Bentley, D. R. 1988. Structural superfamilies of the complement system. *Exp. Clin. Immunogenet. 5:69.*