Coelomocytes Express SpBf, a Homologue of Factor B, the Second Component in the Sea Urchin Complement System

L. Courtney Smith, Chi-Schin Shih and Sebastian G. Dachenhausen

*J Immunol* 1998; 161:6784-6793; 
http://www.jimmunol.org/content/161/12/6784

References
This article cites 54 articles, 21 of which you can access for free at: http://www.jimmunol.org/content/161/12/6784.full#ref-list-1

Subscription
Information about subscribing to *The Journal of Immunology* is online at: http://jimmunol.org/subscription

Permissions
Submit copyright permission requests at: http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts
Receive free email-alerts when new articles cite this article. Sign up at: http://jimmunol.org/alerts
Coelomocytes Express SpBf, a Homologue of Factor B, the Second Component in the Sea Urchin Complement System\textsuperscript{1,2}

L. Courtney Smith,\textsuperscript{3} Chi-Schin Shih,\textsuperscript{4} and Sebastian G. Dachenhausen\textsuperscript{5}

A homologue of factor B, SpBf, has been cloned and sequenced from an LPS-activated coelomocyte cDNA library from the purple sea urchin, Strongylocentrotus purpuratus. The deduced amino acid sequence and domain structure show significant similarity to the vertebrate Bf/C2 family proteins. SpBf is a mosaic protein, composed of five short consensus repeats, a von Willebrand Factor domain, and a serine protease domain. It has a deduced molecular mass of 91 kDa, with a conserved cleavage site for a putative factor D protease. It has ten consensus recognition sites for N-linked glycosylation. Amino acids involved in both Mg\textsuperscript{2+} binding and in serine protease activity in the vertebrate C2/Bf proteins are conserved in SpBf. Phylogenetic analysis of SpBf indicates that it is the most ancient member of the vertebrate Bf/C2 family. Additional phylogenetic analysis of the SCRs indicates that five SCRs in SpBf may be ancestral to three SCRs, which is the typical pattern in the vertebrate Bf/C2 proteins. RNA gel blots show that SpBf transcripts are 5.5 kb and are specifically expressed in coelomocytes. Genome blots suggest that the SpBf gene (Sp152) is single copy gene per haploid genome. This is the second complement component to be identified from the sea urchin, and, with the sea urchin C3 homologue, these two components may be part of a simple complement system that is homologous to the alternative pathway in higher vertebrates.


Echinoderms have a non-specific, non-adaptive immune response that shows similarities to higher vertebrate innate immunity. This has been demonstrated with a number of experimental approaches. Sea stars and sea cucumbers were first used to establish that echinoderms are capable of differentiating between self and non-self tissues, in that they can reject allogeneic skin grafts but would not reject autografts (1, 2). However, immune memory could not be demonstrated because clearance rates of bacteria, xenogeneic cells, and bacteriophage from the coelomic cavities of sea urchins and sea stars did not accelerate with multiple injections (3–5). Furthermore, recognition specificity could not be demonstrated (6) because rejection rates of second set and third party allografts in the sea urchin, Lytechinus pictus, were identical even though they were accelerated relative to the primary rejections (7, 8). Together, these reports defined immunity in echinoderms as a non-adaptive or innate system functioning in the absence of adaptive and specific non-self recognition capabilities typical of vertebrates.

At the molecular level, quantitation of profilin transcripts in coelomocytes was used to characterize the sea urchin immune response as extremely sensitive to minimal injury (9) and to injections of small amounts of LPS (10). Profilin is a key actin regulatory protein that couples signal transduction with cytoskeletal alterations (11). Increases in profilin transcripts in coelomocytes were interpreted to imply changes in cell shape that, in turn, indicated the activation of these amoeboid, phagocytic cells responding to injury or to LPS (9, 10). A subset of sea urchin coelomocytes, the phagocytes, are known for their dramatic cytoskeletal shape changes in response to minor perturbations (12). Although these reports helped to characterize immune reactivity in this species, they did not indicate that the sea urchin system was anything other than a typical, albeit very sensitive, invertebrate immune system (for review, see Ref. 13).

Homology of the innate immune response within the deuterostome lineage of animals that includes the echinoderm (sea urchins) and chordate (vertebrates) phyla was established with the identification of an expressed sequence tag (EST)\textsuperscript{6} 064 from an LPS-activated sea urchin coelomocyte cDNA library (14). GenBank search results indicated that EST064 encoded a new member of the thioester family of complement proteins (15). This conclusion was based on sequence similarities, overall protein structure, and phylogenetic analysis. Identification of a single complement system as a part of the sea urchin immune response established that echinoderms and, by inference, all deuterostome invertebrates share innate immune system homologies with vertebrates. Furthermore, characterizing the simpler immune response exhibited by sea urchins is important for understanding the ancestral deuterostome defense system and for reconstructing the evolutionary changes that occurred during the process of assembling the higher vertebrate immune system.

\textsuperscript{1}This work was supported by National Science Foundation (NSF) Grants MCB9219330, MCB9596251, and MCB9603086 to L.S.C. and by an NSF Research Experience for Undergraduates Award to C.S.S.

\textsuperscript{2}GenBank accession number for the sea urchin factor B sequence is AF059284.

\textsuperscript{3}Address correspondence and reprint requests to Dr. L. Courtney Smith, Department of Biological Sciences, 334 Lerner Hall, George Washington University, 2023 G St. NW, Washington, DC 20052. E-mail address: csmith@gwu.edu

\textsuperscript{4}Current address: School of Medicine, George Washington University, Washington, DC 20052.

\textsuperscript{5}Current address: Naval Nuclear Power Propulsion Program, United States Navy.

\textsuperscript{6}Abbreviations used in this paper: EST, expressed sequence tag; Bf, factor B; SCR, short consensus repeats; vWF, von Willebrand Factor; UT, untranslated; nt, nucleotide; ORF, open reading frame; MASP, mannose-binding lectin-associated serine protease; MBL, mannose-binding lectin.

\textsuperscript{7}In our report of the SpC3 sequence (see Ref. 15), figure 2 incorrectly showed the histidine in position 1090 as being putatively associated with the thioester. The correct histidine is in position 1145.
The suggestion that the alternative complement cascade was the foundation on which aspects of higher vertebrate adaptive immunity that are dependent on complement effector functions may have been built and expanded (14) is reinforced with the characterization of a second sea urchin EST. We report here the complete sequence of EST152, hereafter called Sp152, which encodes a homologue of vertebrate factor B (Bf), called SpBf, which is a new member of the Bf/C2 protein family. In the alternative pathway in higher vertebrates, Bf is the second component to function in the cascade and binds to activated C3b and C3(H2O) (for a review of the alternative complement cascade, see Ref. 16). The SpBf domain structure is typical of the Bf/C2 complement family of proteins in vertebrates. It is a mosaic protein composed of five short consensus repeats (SCRs) (which are sometimes referred to as complement control protein (CCP) modules), a von Willebrand Factor (vWF) domain, and a serine protease domain. Alignments with other Bf/C2 proteins show that SpBf has conserved amino acids for binding Mg2+ and a conserved cleavage site for a putative factor D. Results from phylogenetic analyses indicate that SpBf is the most ancient Bf/C2 family member and that five SCRs may be the ancestral condition rather than three SCRs, which is typical for the vertebrate proteins. Transcripts from Sp152 are specifically expressed in coelomocytes and appear to be generated from a single copy gene. The two sea urchin proteins, SpC3 and SpBf, appear to be homologous to the two-component complement system that has osprexin functions in anagnaths (17–20) and to the alternative cascade in higher vertebrates. We hypothesize that these two proteins act together and that this sea urchin complement system also functions to opsonize foreign cells and particles, augmenting their phagocytosis and subsequent destruction by the coelomocytes.

Materials and Methods

RNA isolation

Total RNA was isolated from coelomocytes and other adult tissues as previously described (9, 15). Briefly, coelomic fluid (40 ml) was poured through sterile cheese cloth and mixed into 10 ml of cold Ca2+ and Mg2+-free sea water (21) containing 30 mM EDTA, pH 7.4 (CMFSW-E). Coelomocyte pellets and minced solid tissues were vortexed and homogenized using a dounce homogenizer in guanidinium thiocyanate extraction buffer (5 M guanidinium thiocyanate, 50 mM NaOAc, 50 mM EDTA, 50 mM Tris (pH 7.4), 5% 2-ME) to which was added N-lauroyl sarcosine to a final concentration of 2%. Total RNA was pelleted through a cushion of 5.7 M CsCl containing 50 mM NaOAc and 50 mM EDTA at 10 000 × g in either a Ti60 fixed angle rotor (Beckman Instruments, Fullerton, CA) or a swinging bucket rotor (Sorvall, Newtown, CT) at 20°C for 20 h. Pellets were washed in 70% ethanol, resuspended in RNase-free water, extracted in (1:1) phenol/chloroform/isoamyl alcohol, precipitated, and resuspended in RNase-free water. Poly(A)+ RNA was isolated using oligo(dT) magnetic beads (Dynal, Great Neck, NY) according to the manufacturer’s instructions.

cDNA library construction

The sea urchin immune response was activated by injections of LPS, and activated coelomocyte RNA was isolated (10). An activated coelomocyte cDNA library was constructed from poly(A)+ RNA using the Time-Saver cDNA kit and directionally cloned into the AExCell phage (Pharmacia, Piscataway, NJ) as previously described (14, 15). The library was screened using 32P-labeled RNA probes that were generated according to technical information from Promega (Madison, WI) and as previously reported (9, 15).

Sequencing

DNA sequencing was conducted on plasmid DNA according to the dioxygenucleotide termination protocol (22) using the TaqTrack sequencing kit (Promega) incorporating [α-32P]dATP (DuPont/NEN, Boston, MA). Sequencing reactions were electrophoresed on a 6% acrylamide gel with 0.6× TBE (10× TBE is 0.9 M Tris, 0.9 M Boric acid, 20 mM EDTA, pH 8.3) running buffer, after which the gel was dried and exposed overnight to BioMax MR-1 x-ray film (Eastman Kodak, Rochester, NY). For DNA sequencing was conducted on plasmid DNA according to the dioxygenucleotide termination protocol (22) using the TaqTrack sequencing kit (Promega) incorporating [α-32P]dATP (DuPont/NEN, Boston, MA). Sequencing reactions were electrophoresed on a 6% acrylamide gel with 0.6× TBE (10× TBE is 0.9 M Tris, 0.9 M Boric acid, 20 mM EDTA, pH 8.3) running buffer, after which the gel was dried and exposed overnight to BioMax MR-1 x-ray film (Eastman Kodak, Rochester, NY). For clones 152L and 152-69X2, which were longer than several hundred nucleotides (see Fig. 1), inserts were subcloned into the Bluescript vector (Stratagene, La Jolla, CA), and the Erase-a-Base kit (Promega) was employed to create a nested set of deleted clones. Sets of overlapping insert deletions were sized by PCR, and sequences were assembled using the DNASIS sequence analysis program (Hitachi Software, San Francisco, CA) on a pentium personal computer.

RNA gel blots

Poly(A)+ RNA (0.4 µg), isolated with oligo(dT) magnetic beads (Dynal), was electrophoresed through a 0.8% agarose gel containing 2.2 M formaldehyde in 1× MOPS buffer (20 mM 3-N-morpholino)propanesulfonic acid, 5 mM NaOAc, 1 mM EDTA, pH 7) and capillary blotted onto GeneScreen Plus (DuPont/NEN). The blot was probed with a 32P-labeled PCR-amplified DNA insert clone 152S (500 bp; Fig. 1) or the control clone EST219 (900 bp), which encodes a homologue of human L8 ribosomal protein (14). The PCR reaction was performed on a 9600 PCR machine (Perkin-Elmer, Norwalk, CT) and contained 200 ng template, 1.0 µM each primer (Sp6 or T3 and T7), 7.5 µM dNTPs, 2.5 mM MgSO4, 10 µCi [a-32P]dCTP (DuPont/NEN), and 1 U Taq polymerase (Promega), with buffer supplied by the company. The thermocycler was programmed as follows: 94°C for 5 min followed by 20 to 30 cycles of 30 sec at 94°C, 30 sec at 57°C, 1 min at 72°C, finishing with 72°C for 2 min. The entire PCR reaction was passed through a G-50 Sephadex (Pharmacia) spin column to remove unincorporated nucleotides. The probe was denatured at 100°C for 2 min before being added to the hybridization solution.

Filters were prehybridized for 2 h in hybridization solution (50% formamide, 5× saline sodium phosphate buffer [50 mM NaOAc, 1× EDTA, pH 7.4, 1× BSA, 7% SDS) and then hybridized with the probe at 42°C overnight in a rotating oven (Robbins Scientific, Sunnyvale, CA). Final washes were conducted at 68°C in 1× SSC (0.15 M NaCl, 15 mM Na Citrate, pH 7) with 1% SDS. Filters were exposed overnight to X-Omat XAR-5 x-ray film (Eastman Kodak). Transcript sizes were estimated from RNA standards (Bio-Rad, Hercules, CA). Reprobing was conducted after the blots were stripped in 0.1× SSC at 100°C for 15 min.

Protein alignments and phylogenetic analysis

A basic BLAST search of GenBank (23) was done using the deduced amino acid sequence of SpBf to identify sequence matches to other proteins. The BLAST list included the Bf/C2 protein family members in addition to other mosaic proteins containing vWF domains, serine protease domains, and SCRs. All of the Bf/C2 members and several additional matched sequences were used to construct protein alignments with the CLUSTAL W program, using default parameters (24).

To identify phylogenetic relationships among the sea urchin and vertebrate Bf/C2 proteins, sequences were first aligned with the CLUSTAL W program (24) and were then imported into the PAUP program (version 3.1.1) (25). Outgroups were identified from BLAST results, and the heuristic search method was used to obtain the shortest tree. The heuristic search in PAUP was set for tree-bisection-reconnection branch-swapping with an initial MAXTREES setting of 100, with all data weighted equally. The general search options were set to keep minimal trees only and to collapse zero-length branches. When multiple trees were obtained, a strict consensus tree was calculated. A number of analyses were done on full-length sequences of the Bf/C2 proteins, on SCR-deleted sequences, and on independent SCR sequences. Different outgroups were chosen to root individual trees, which were based on either sequence similarities to the vWF domain or the serine protease domain, or were SCRs with known binding function or lack thereof. In some SCR analyses, additional SCRs with known binding function were added to the Bf/C2 ingroup.

Results

Isolation and sequencing Sp152

One of the partially sequenced cDNAs that was reported as EST152 from the purple sea urchin matched to SCR domains from complement receptors and regulatory proteins (14). Although our analysis indicated that the EST152 BLAST matches were below significance (see Q value, Table II in Ref. 14), matches were mostly restricted to the consensus amino acids in the SCR domains. Consequently, we identified two complete and two partial SCRs in the EST152 protein sequence (see Fig. 2 in Ref. 14). However, since the sequence of the EST152 clone began within the ORF, the library was rescreened with a riboprobe made from the 152S subclone (Fig. 1). We picked 74 positives, which were
From the 5'-end there are four start-translation codons (ATGs), but, because the clone with the longest amplified band, pExCell152-69, was chosen for further characterization (Fig. 1). The 5'-end of pExCell152-69 (152-69X2) was subcloned into Bluescript (Stratagene) and used as a template to produce an antisense riboprobe to rescreen the library and obtain the second clone, pExCell152-69, to screen the RNA gel blot and to screen the genome blot.

An SCR domain is typically about 60 amino acids long with a number of conserved residues. These amino acids include four cysteines, three glycines, two prolines, two tyrosines (or phenylalanine), and one tryptophan, all of which are located at specific positions within the domain (29). These consensus positions are shown in Fig. 3. SpBf has one cysteine located in the vWF domain and corresponds to cleavage sites in the last cysteine in the last SCR for each protein. The alignment reveals a number of highly conserved amino acids. These include a factor D cleavage site at the beginning of the vWF domain, the five amino acids that are involved in binding Mg^{2+} (28), and the amino acids involved in the serine protease activity (Fig. 3).

There are 30 cysteines in SpBf, 20 of which are located in the SCR domains (Fig. 4), and the remaining 10 are shown in the alignment (Fig. 3). SpBf has one cysteine located in the vWF domain, which is a region in the vertebrate Bf/C2 proteins where no cysteines are found. In the serine protease domain, SpBf has nine cysteines. Eight align perfectly with cysteines in the Bf/C2 proteins (Fig. 3) and with cysteines in other serine proteases chosen from the BLAST results (data not shown). In addition, there is one cysteine in SpBf that does not align with the vertebrate Bf/C2 proteins, and there are two positions where cysteines align in all the vertebrate proteins but are missing in SpBf (Fig. 3).

An SCR domain is typically about 60 amino acids long with a number of conserved residues. These amino acids include four cysteines, three glycines, two prolines, two tyrosines (or phenylalanine), and one tryptophan, all of which are located at specific positions within the domain (29). These consensus positions are shown in Fig. 4 where the five SCR domains from SpBf are aligned to each other. Two disulfide bonds are formed between the four cysteines and maintain the topology of the domain (Fig. 4). Although some of the consensus amino acids are missing from SpBf, SCR 4 (two tyrosines) and SCR 5 (the third glycine), all four cysteines are present in each SCR, suggesting that these domains fold as expected. In our previous report (14), we suggested that one SCR was missing a cysteine; however, this was due to sequencing errors.

Phylogenetic relationship between SpBf and other Bf/C2 family proteins

Since the protein encoded by Sp152 appears to be a new member of the Bf/C2 family of proteins, we were interested to know whether it was more similar to Bf or C2. Pairwise alignments between SpBf and all the other Bf/C2 proteins, with and without
FIGURE 2. cDNA sequence and deduced protein sequence of SpBf. The sequence was obtained from the two overlapping cDNA clones and the set of nested deletion clones that were obtained from 152L and 152-69X2 using the sequence alignment program from DNASIS (Hitachi Software) on a pentium personal computer. The deduced amino acid sequence is shown below the DNA sequence. These sequences have been deposited in GenBank with the accession number AF059284. Stop codons are underlined in both UT regions. Four ATGs are shown in bold in the 5′ UT region. The start site is not identified by a Kozak sequence; however, a hydrophobic region, the putative leader, is underlined at the N terminus of the protein. The domains are labeled in the figure including the consensus positions for serine protease activity and a conserved factor D cleavage site. The consensus N-linked glycosylation sites are noted with CHO. One AU-rich element, ATTTA, is denoted in bold and is present in the 3′ UT region.

The Journal of Immunology
the SCR domains, were used to calculate percentage of amino acid similarities and identities between the proteins (Table I). Results of this analysis show that SpBf is about equally similar to all Bf/C2 protein family members. Differences in the number of charged amino acids in exon 15 from human Bf and C2 genes have been suggested as a means to differentiate between these two genes (30), and this approach has been used to characterize the Bf/C2 homologue from the medaka fish (31). There are very few charged amino acids in the region of SpBf that align with the diagnostic exon 15, but this may have been due to the overall similarity between SpBf and human Bf and C2 in this region being very poor (Fig. 3). Consequently, to assess relationships among the Bf/C2 protein sequences, alignments of full length and SCR-deleted sequences of the sea urchin and vertebrate proteins were done for phylogenetic analysis. Outgroups consisted of either the three A domains from vWF or three serine proteases chosen from the BLAST results. All trees generated by these methods were similar, and a representative tree, using full-length sequences and the three vWF A domains as the outgroup, is shown in Fig. 5. The alignment that was used for this analysis is available by e-mail (see legend to Fig. 5). Although some bootstrap numbers are low for some branches, which correspond to minor differences between trees, the position of SpBf at the base of the Bf/C2 clade was consistent in all trees. This suggests that SpBf predates the Bf/C2 duplication

FIGURE 3. Alignment of the vWF and serine protease domains from SpBf and the vertebrate Bf/C2 family proteins. To simplify the analysis, the SCR domains plus five amino acids were deleted, and the remaining domains were aligned with CLUSTAL W program (24), with some minor adjustments by hand. Positions of identity and similarity are identified below the alignment with stars and dots respectively. Domains and amino acids with functional significance are labeled. Location of exon 15 from HsBf was identified from Volanakis (16). Accession numbers for sequences used in this alignment can be found in the legend for Fig. 5. Bf, factor B; C2, complement component C2; Sp, Strongylocentrotus purpuratus; Lj, Lampetra japonica, lamprey; Br, Brachydanio rerio, zebra fish; Ol, Oryzias latipes, medaka fish; Xl, Xenopus laevis (A gene); Mm, Mus musculus; Hs, Homo sapien.

FIGURE 4. Alignment of the SpBf SCRs. The five SCRs from SpBf are aligned to each other with the consensus amino acids shown in bold and at the bottom. Two disulfide bonds maintain the SCR domain structure and are formed between the first and third cysteine, and the second and fourth cysteine. The alignment was done with CLUSTAL W (24) with significant adjustments by hand.
event that appears to have occurred at some point during the evolution of vertebrates (32) and, therefore, should be considered a Bf homologue. Furthermore, since the thioester-containing complement component that has been identified in the sea urchin is a homologue of complement C3, being less similar to C4 (15), this implies the presence of a simple alternative pathway and, in turn, this also indicates that SpBf is a homologue of vertebrate Bf.

Although SpBf shows significant sequence similarities to the Bf/C2 family proteins, it was important to determine whether the five SCR domains were a result of domain duplications in echinoderms. To understand the SpBf protein structure in more detail, sequence similarities (or differences) between the five SCR domains from SpBf and the three SCRs from the vertebrate Bf/C2 proteins (i.e., the ingroup) were assessed through amino acid alignments and phylogenetic analyses. The SCR domains were employed as independent sequences beginning with the first consensus cysteine and ending with the linker region between the SCR domains (four to seven amino acids). Because the SCR domain is short, about 60 amino acids, the degree of support for the branches within an individual tree was low. Consequently, a large number of trees were generated from 18 runs that employed additional SCRs from other SCR-containing proteins. These additional SCRs were either used in the outgroup or were added to the ingroup (Table II). For the 18 phylogenetic analyses that were done, the choice of SCRs for the outgroup was based on a) distant phylogenetic relationship to deuterostomes (Table II, run no. 18; Ref. 33), b) close phylogenetic relationship to deuterostomes (Table II, run nos. 15–17; Refs. 34–38), c) SCRs with known protein binding function (Table II, run nos. 7–14; Refs. 39–42), d) SCRs with putative spacer function (Table II, run nos. 1–3; Ref. 43), and e) SCRs with known three-dimensional structure (Table II, run nos. 4–6; Refs. 44 and 45). In some cases, SCRs with documented C3 binding function were included with the ingroup (Table II, run nos. 2–6) to increase the ingroup size. The number of trees generated from individual runs ranged from 1 to 20 (Table II), and, when more than one tree was obtained, strict consensus trees were calculated for subsequent analysis.

All trees were inspected, and the frequencies with which the vertebrate SCRs clustered into independent clades are shown in Table IV and the frequencies with which SpBf SCRs clustered with vertebrate SCR clades are shown in Table III. Because no individual tree demonstrated all the results shown in Tables III and IV, no tree is shown. This analysis revealed several interesting points. First, the three SCRs from the vertebrate Bf/C2 proteins tend to cluster into separate clades rather than to form multiple or mixed clades, indicating that this approach for amino acid sequence comparisons is sensitive enough to identify sequence differences between the SCRs in the Bf/C2 proteins (Table IV). Second, the five SpBf SCRs tend to cluster with certain vertebrate SCR clades (Table III). This result suggests that a) SpBf SCRs 1 and 2 are most similar to vertebrate SCR 1, b) SpBf SCR 3 is most similar to vertebrate SCR 2, c) SpBf SCR 4 is similar to both vertebrate SCRs 2 and 3, and d) SpBf SCR 5 is most similar to vertebrate SCR 3. In addition, in almost all cases, when a sea urchin SCR clustered with a vertebrate SCR clade, it almost always fell at the base of the clade. The one exception was that, when SpBf SCR 5 clustered with vertebrate SCR 3 clade, in three of twelve cases it was positioned terminally (data not shown). In general, this analysis shows that the SCRs in SpBf are in the same relative order in the protein as the SCRs in the vertebrate Bf/C2 proteins in terms of sequence similarity. Furthermore, this analysis indicates that a structural condition of five SCRs may be ancestral for the Bf/C2 protein family and that, therefore, the vertebrate proteins may have lost two. The information presented here on phylogenetic relationships that includes sequence similarities between the SCRs (Tables II, III, and IV) and the overall comparisons of the SpBf and the vertebrate Bf/C2 proteins (Fig. 5) suggest

### Table I. Amino acid comparisons between SpBF and other Bf/C2 family proteins

<table>
<thead>
<tr>
<th>Vertebrate Bf/C2 Proteins</th>
<th>Full-Length Sequences</th>
<th>Sequences Excluding SCRs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% Identical</td>
<td>% Similar plus identical</td>
</tr>
<tr>
<td>LjBf</td>
<td>24.1</td>
<td>36.7</td>
</tr>
<tr>
<td>BrBf</td>
<td>20.4</td>
<td>35.9</td>
</tr>
<tr>
<td>OIBf/C2</td>
<td>19.9</td>
<td>32.9</td>
</tr>
<tr>
<td>XIBf/B</td>
<td>22.6</td>
<td>34.1</td>
</tr>
<tr>
<td>MnBf</td>
<td>22.5</td>
<td>36.1</td>
</tr>
<tr>
<td>HsBf</td>
<td>21.3</td>
<td>35.6</td>
</tr>
<tr>
<td>MnC2</td>
<td>22.7</td>
<td>37.7</td>
</tr>
<tr>
<td>HsC2</td>
<td>23.2</td>
<td>36.5</td>
</tr>
</tbody>
</table>


### Figure 5. Phylogenetic relationships between members of the Bf/C2 family proteins.

The tree was constructed using the PAUP program (25) from an alignment done with CLUSTAL W (24). The alignment used to generate this tree can be obtained by contacting the EMBL server by e-mail. Send the request to Netserv@ebi.ac.uk and include in the message GET ALIGN:DS33817.DAT. Two analyses were done using different outgroups. The figure shows the result when three A domains from the human vWF protein were used as the outgroup. The outgroup for the second tree included three serine proteases that were chosen from the BLAST search results (tree not shown). No significant differences between these two trees were seen with respect to the position of SpBf within the vertebrate Bf/C2 clade. Minimum tree length was 3107. The degree of support for internal branches was assessed using the bootstrapping method with 1000 replications, and bootstrap numbers are shown. Accession numbers: SpBf, AF059284; LjBf, D13568; BrBf, U34662; OIBf, D84063; XIBf/A, D29796; XIBf/B, D49373; MnBf, M57890; MnC2, M57891; HsBf, X72875; HsC2, X04481; HsVWF, X04385.
that SpBf should be considered a new and ancestral member of the Bf/C2 protein family.

Sp152 gene expression in sea urchin tissues

We have previously shown that transcripts from the SpC3 gene, Sp064, are specifically expressed in coelomocytes (15). Since the original clone, EST152, was obtained from the an LPS-activated coelomocyte cDNA library (14), we were interested to know whether expression of the Sp152 gene was also coelomocyte specific. The poly(A) RNA gel blot of major sea urchin tissues (coelomocytes, ovary, testis, and gut), originally probed for Sp064 transcripts (see Fig. 6 in Ref. 15), was stripped and reanalyzed with a PCR probe generated from the 152S subclone (Fig. 1). The blot shows that a single transcript of 5.5 kb is present in coelomocytes and that this band is either very weak or absent in the other major adult tissues (Fig. 6). The weak bands seen in the gonad tissues may be due to low level expression in these sea urchin tissues or may be due to expression by coelomocytes that were present in or on these organs at the time of tissue collection and RNA isolation.

Table II. SCRs used as outgroups and additions to the ingroup for amino acid sequence analysis of SCRs from SpBf and the vertebrate Bf/C2 family

<table>
<thead>
<tr>
<th>Run No.</th>
<th>SCR Source</th>
<th>SCRs Used as the Outgroup</th>
<th>SCRs Added to the Ingroup</th>
<th>Trees Generated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HsFactorH 15th</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>HsFactorH 15th</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>HsFactorH 15th</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>HsFactorH 16th</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>HsFactorH 16th</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>HsFactorH 16th</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>HsFactorH 1st</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>HsFactorH 2nd</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>HsFactorH 3rd</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>HsFactorH 2nd and 3rd</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>HsDAF 1st</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>HsDAF 2nd</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>HsDAF 3rd</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>HsDAF 4th</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>HmMASPb 2 SCRs each</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HmMASPb</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HmMASP1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HmMASP2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>HsC1r &amp; HsCls</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>HmMASPb &amp; b 2 SCRs each</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HmMASP1 &amp; 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HsC1r and Cls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>TtFactorC</td>
<td>5 SCRs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The ingroup consisted of the five SCRs from SpBf and the three SCRs from each of the vertebrate Bf/C2 proteins shown in Table I. Protein alignments were done on CLUSTAL W (24), and phylogenetic relationships were analyzed with the PAUP program (25). DAF, decay-accelerating factor; Hr, Halocynthia roretzi, tunicate; Tt, Tachypleus tridentatus, horseshoe crab. Accession numbers: HsFactorH, Y00716; HsDAF, M30142; HrMASPa, D88204; HrMASPb, D88205; HsMASP1, D28593; HsMASP2, Y09926; HsC1r, M14058; HsCl, X06596; TtFactorC, D90271.

Table III. Frequency of SpBf SCRs to cluster with certain vertebrate Bf/C2 SCRs

<table>
<thead>
<tr>
<th>SpBF SCRs</th>
<th>SCR 1</th>
<th>SCR 2</th>
<th>SCR 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCR 1</td>
<td>67</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>SCR 2</td>
<td>60</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>SCR 3</td>
<td>20</td>
<td>50</td>
<td>30</td>
</tr>
<tr>
<td>SCR 4</td>
<td>13</td>
<td>48</td>
<td>43</td>
</tr>
<tr>
<td>SCR 5</td>
<td>16</td>
<td>32</td>
<td>51</td>
</tr>
</tbody>
</table>

* The clustering of the SpBf SCRs with vertebrate SCR clades was assessed from 18 phylogenetic analyses using outgroups with some additions to the ingroup as shown in Table II.

Table IV. Frequency of SCRs from vertebrate Bf/C2 proteins to cluster in single vs multiple clades

<table>
<thead>
<tr>
<th>Vertebrate Bf/C2 SCRs</th>
<th>Clusters in a Single Clade (%)</th>
<th>Clusters in Multiple Clades (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCR 1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>SCR 2</td>
<td>83</td>
<td>17</td>
</tr>
<tr>
<td>SCR 3</td>
<td>67</td>
<td>33</td>
</tr>
</tbody>
</table>

* The formation of a single or multiple clades of vertebrate SCRs was assessed from 18 phylogenetic analyses using various outgroups and some additions to the ingroup as shown in Table II.

Sp152 gene copy number

In most vertebrates, Bf is a single copy gene, although Bf in Xenopus and trout appears recently duplicated (46, 47) and Bf and C2 are generally considered to have arisen from an ancient duplication in a common ancestor of mammals (48). To determine whether Sp152 is a single copy gene per haploid genome, we probed a genome blot of sea urchin DNA isolated from three individuals and digested as previously described (15, 49). The genome blot was analyzed at high stringency with a PCR-generated DNA probe that corresponded to the SCR region of the ORF (500 bp, 152S; see Fig. 1). One or two large bands ($\geq12$ kb) were seen for all three individuals when the DNA was digested with KpnI or BamHI, suggesting that Sp152 is a single copy gene (data not shown).
However, the EcoRI digests in all lanes showed a pattern of eight smaller bands ranging in size from 0.5 to 8 kb. This result is more difficult to understand without data on the intron/exon structure of the Sp152 gene, but at present our interpretation of the multiple bands in the EcoRI digest for the three sea urchins is as follows. The probe sequence spans the region of the cDNA from nt 522 to nt 1032 (Fig. 2), which includes part of two and all of two more SCRs. Because these domains are generally known to be contained within separate exons (50, 51), this suggests that the probe may bind to four exons in the Sp152 gene. If each of the three introns between the SCR exons contains an EcoRI site, this would result in eight bands on an EcoRI digest for a diploid animal. In general, the genome blot results are consistent with 1) KpnI and BamHI sites located outside of the SCR regions from two alleles resulting in one or two large fragments hybridizing with the probe, and 2) EcoRI sites located between each of the SCR exons, which would separate them into smaller fragments to which the probe would hybridize. These results, as interpreted, are consistent with a single copy gene per haploid genome.

Discussion

This is the first identification of a factor B homologue from an invertebrate. Like vertebrate Bf/C2 proteins, SpBf is a mosaic protein composed of SCRs, a vWF domain, and a serine protease domain. Highly conserved regions in the SpBf protein include the SCR consensus positions, a putative factor D cleavage site and Mg\(^{2+}\) binding sites within the vWF domain, and conserved positions within the serine protease domain required for protease activity (Figs. 2 and 3). The homology to Bf rather than C2 is based on sequence comparisons from alignments (Fig. 4), phylogenetic trees (Fig. 5), and BLAST searches (lamprey Bf was consistently the best match, with the top 12 matches being to other Bf or C2 proteins, data not shown). The Bf homology is also supported by the fact that C2 proteins have not been found in animals other than mammals (52–54), indicating that, in nonmammalian vertebrates, Bf must function in both classical and alternative pathways (32). This has, in fact, been shown to occur in trout where two Bf proteins, these SpBf SCRs may be more important in protein interactions (47). Similarly, for human decay-accelerating factor, SCRs two through four are essential for binding C3b and inhibiting the alternative pathway (41), with the third being most important and the first being least important (40). Together, these studies show that the first SCR in the vertebrate proteins is not very important for binding function, while the third SCR is essential. These functional analyses of the three vertebrate Bf/C2 SCRs and their respective importance in complement binding is reflected in our phylogenetic analyses by the clustering of these domains into independent clades (Table IV).

Phylogenetic analyses were initially used to identify sequence similarities between sea urchin and vertebrate SCRs, but, because our results show that the three vertebrate SCRs cluster into separate clades (Table IV), we have assumed that sequence similarity can be used to infer functional similarity. Consequently, our data suggest that, since SCRs 1 and 2 from SpBf are similar to SCR 1 from the vertebrate Bf/C2 proteins, these domains may be less important in protein interactions. Similarly, since SCRs 3, 4, and 5 from SpBf are more similar to SCRs 2 and 3 from the vertebrate proteins, these SpBf SCRs may be more important in protein interactions.

Model of the sea urchin complement cascade

We hypothesize that the SpBf and SpC3 proteins function together as part of a simple complement system to opsonize foreign cells, particles, and molecules that augments their removal and degradation by phagocytic coelomocytes. This simple sea urchin complement pathway might function like the “archeo-complement system” that was first proposed by Lachmann (59), which would make it essentially homologous to the alternative pathway in vertebrates. Opsonization would begin with spontaneously activated SpC3 in solution, in the form of SpC3(H\(_2\)O), which would become bound as SpC3b to a foreign surface by its thioester site and would subsequently be bound by SpBf. The change in conformation of bound SpBf would result in its cleavage by a putative factor D. The possible interaction between SpC3 and SpBf is supported by the conserved Mg\(^{2+}\) binding sites and the conserved cleavage site for factor D in the SpBf sequence (Figs. 2 and 3). The involvement of Mg\(^{2+}\) and the Mg\(^{2+}\) binding sites within the vWF domain has been implicated in Bb binding to C3b after factor D cleavage of vertebrate Bf (28). The SpC3b-SpBb complex would then function as a C3 convertase through the activation of the SpBf serine protease domain, creating an amplification feedback loop to cleave and activate more SpC3 for deposition onto the foreign surface. The feedback loop is inferred from the conserved C3 convertase cleavage site in the deduced SpC3 sequence (see Fig. 3 in Ref. 15) and from the sizes of SpC3 fragments in activated coelomic fluid (W. Al-Sharif and L. C. Smith, unpublished observations). Fragments indicate that SpC3 convertase cleavage site is functional (W. Al-Sharif and L. C. Smith, unpublished observations). The feedback loop of the archeo-complement system would result in efficient opsonization of foreign cells or particles that, in turn, would act to augment phagocytosis by coelomocytes bearing a putative complement receptor. The efficiency, speed, and extent of opsonization that occurs as a result of this feedback loop would be advantageous over simple opsonins and might make this simple complement system an important defense mechanism for the sea urchin.
Acknowledgments

We acknowledge the support and encouragement from Eric Davidson, in whose laboratory this work was initiated. We also express our gratitude to W. Al-Sharif for his time and assistance during the course of this work. M. Allard, J. Miller, M. Daly, and P. Herendeen gave their time and advice. O. Sunyer for significant input on the phylogenetic analyses. We thank M. Allard, J. Miller, M. Daly, and P. Herendeen gave their time and advice. O. Sunyer for significant input on the phylogenetic analyses. We thank M. Allard, J. Miller, M. Daly, and P. Herendeen gave their time and advice.

Immunology within the deuterostomes

The echinoderm immune system has previously been characterized as a nonspecific, nonadaptive response (6, 8), and there was nothing about this system that characterized it as anything other than a typical invertebrate immune system. However, this has now changed. With the identification of a simple complement system in the sea urchin, this establishes immune homology within the entire lineage of the deuterostomes and differentiates immune functions in deuterostome invertebrates from those in protostomes. Continued analysis of the sea urchin complement system not only will clarify our characterization of the echinoderm immune response but also, through comparisons with homologous systems in the vertebrates, will advance our understanding of changes that have occurred during evolution in the deuterostome lineage that culminated in the complex and multi-tiered immune system that functions in mammals.

References