Developmental and Comparative Immunology 140 (2023) 104584

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

journal homepage: www.elsevier.com/locate/devcompimm

# The echinoid complement system inferred from genome sequence searches

L. Courtney Smith [a,*], Ryley S. Crow [a,1], Nicola Franchi [b,1], Catherine S. Schrankel [c,1,2]

[a] Department of Biological Sciences, George Washington University, Washington DC, USA
[b] Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy
[c] Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California, La Jolla, CA, USA

ARTICLE INFO

*Keywords:*
Echinoderm
SMART domains
Thioester proteins
Serine proteases
Lectin pathway
Alternative pathway
Regulatory proteins
Macpf proteins

ABSTRACT

The vertebrate complement cascade is an essential host protection system that functions at the intersection of adaptive and innate immunity. However, it was originally assumed that complement was present only in vertebrates because it was activated by antibodies and functioned with adaptive immunity. Subsequently, the identification of the key component, SpC3, in sea urchins plus a wide range of other invertebrates significantly expanded the concepts of how complement functions. Because there are few reports on the echinoid complement system, an alternative approach to identify complement components in echinoderms is to search the deduced proteins encoded in the genomes. This approach identified known and putative members of the lectin and alternative activation pathways, but members of the terminal pathway are absent. Several types of complement receptors are encoded in the genomes. Complement regulatory proteins composed of complement control protein (CCP) modules are identified that may control the activation pathways and the convertases. Other regulatory proteins without CCP modules are also identified, however regulators of the terminal pathway are absent. The expansion of genes encoding proteins with Macpf domains is noteworthy because this domain is a signature of perforin and proteins in the terminal pathway. The results suggest that the major functions of the echinoid complement system are detection of foreign targets by the proteins that initiate the activation pathways resulting in opsonization by SpC3b fragments to augment phagocytosis and destruction of the foreign targets by the immune cells.

## 1. Introduction and overview of the complement system in mammals

In the late 1800s, immunologists first identified factors in mammalian serum with activity against bacteria, of which one was heat stable and the other was heat sensitive and both were complementary to the functions of the other factor (Ehrlich and Morgenroth, 1899). Based on this report, the name 'complement' was established for the system that is complementary to antibody activity. Although the immunologists assumed only two components at that time, the mammalian complement system is composed of about 50 proteins (Merle et al., 2015a; Sarma and Ward, 2011) that are present in fluid phase and on cell surfaces, of which many are inactive and require proteolytic cleavage for activation. The major outcomes of complement activation are lysis of foreign cells including pathogens, and opsonization of cells and molecules that are targeted as foreign and detected by complement receptors on phagocytes for endocytosis or phagocytosis and degradation. Furthermore, two of the small peptide fragments, the anaphylatoxins, are released upon cleavage of the thioester proteins and function to attract and activate phagocytes thereby augmenting inflammation and amplifying the initial phases of the immune response. Overall, the complement system functions as a surveillance system for defense against pathogens with impacts on inflammation and host homeostasis (Merle et al., 2015b).

The general public tends to believe that antibodies kill pathogens, likely because it is not necessary to describe the complexity of the complement system in explanations of the pathogen killing functions of the immune system. Antibodies function as the activator of the classical pathway (Fig. 1) and link the complement system to the adaptive immune system based on whether antibodies are present for binding to foreign antigens. The next protein in the classical pathway is C1q that is a multimer and binds to the Fc region of the antibody. However, in the

* Corresponding author.
*E-mail address:* csmith@gwu.edu (L.C. Smith).
[1] These authors are listed in alphabetical order and contributed equally to this paper.
[2] Current address, Department of Biology, San Diego State University, San Diego CA, USA.

https://doi.org/10.1016/j.dci.2022.104584
Received 30 June 2022; Received in revised form 1 October 2022; Accepted 31 October 2022
Available online 4 November 2022

absence of antibodies, C1q, functions as a pathogen recognition receptor (PRR; see Table 1 for definitions of abbreviations) and binds several self targets such as other PRRs bound to non self-targets, and also binds directly to non-self targets including a variety of pathogens (Bohlson et al., 2007; Reid, 2018; Thielens et al., 2017). The next step in the classical pathway is the activation of two serine proteases, C1r and C1s, that are intercalated within the C1q multimer. When C1q binds a target that changes the spacial conformation of the multimeric subunits, this auto-activates the protease function of C1r (Gaboriaud et al., 2012). Activated C1r cleaves and activates C1s that, in turn, cleaves and activates C4 and C2. C4 is a thioester protein that, when cleaved to the C4b form, exposes the thioester site that forms a covalent bond with amine and hydroxyl groups on molecules and on cell surfaces. When C2 associates with C4b, it is cleaved by C1s to C2a that remains with C4b on the cell surface and together form the classical pathway C3 convertase (C4b2a) that is capable of cleaving and activating C3. C3 is the key thioester protein and, when cleaved and activated to C3b, also forms covalent bonds with amine and hydroxyl groups on any molecule including those on cell surfaces. The complement regulatory system protects self from this attack such that C3b binds to and accumulates on pathogen surfaces and acts as an opsonin. The smaller fragment, C3a, functions as an anaphylatoxin to induce an inflammatory reaction required for immune responsiveness.

The PRR activities of C1q, which has broad binding capabilities including foreign carbohydrates, can initiate the lectin activation pathway in addition to two other proteins with similar multimeric structures (Reid, 2018). These are mannose binding lectin (MBL) and several ficolins that have a wide variety of binding targets. MBL-associated serine proteinases (MASPs) are intercalated within the MBL multimer (Schwaeble et al., 2002; Wallis, 2007), with similarities to the intercalation of C1r and C1s with the C1q multimer. When MBL is bound to a foreign carbohydrate, the MASPs cleave and activate either C3 or C4 plus C2 that results in the exposure of the thioester sites in C3b or C4b, covalent binding to the foreign surface, and the formation of the C3 convertase complex with cleaved and activated factor B (Bf) or C2, respectively (Fig. 1) (Endo et al., 2011; Fujita, 2002; Fujita et al., 2004b; Matsushita and Fujii, 1992; Takahashi et al., 2006). There are three ficolins in humans that have a similar domain structure as MBL, but rather than a C-type lectin binding domain, they have a fibrinogen domain at the C-terminus, and bind to different carbohydrates than those bound by MBL (Fig. 1) (Liu et al., 2005). The classical and lectin activation pathways are essentially the same, except that the activating proteins differ; either antibody and the C1 complex vs. MBL, ficolins, or C1q.

The third activation pathway, the alternative pathway, does not require recognition of a PAMP or other foreign molecule for initiation, but is based on the spontaneous activation of C3 by a nucleophilic attack from water to form an active C3($H_2O$) that exposes the thioester site in a process called the C3-tickover (Lachmann, 2009). C3($H_2O$) occurs in solution, although it may bind to a surface if the tickover occurs in close proximity to a cell. C3(H2O) can be bound by Bf, which is cleaved by factor D (Df) to activate the serine protease domain in the Bb fragment to

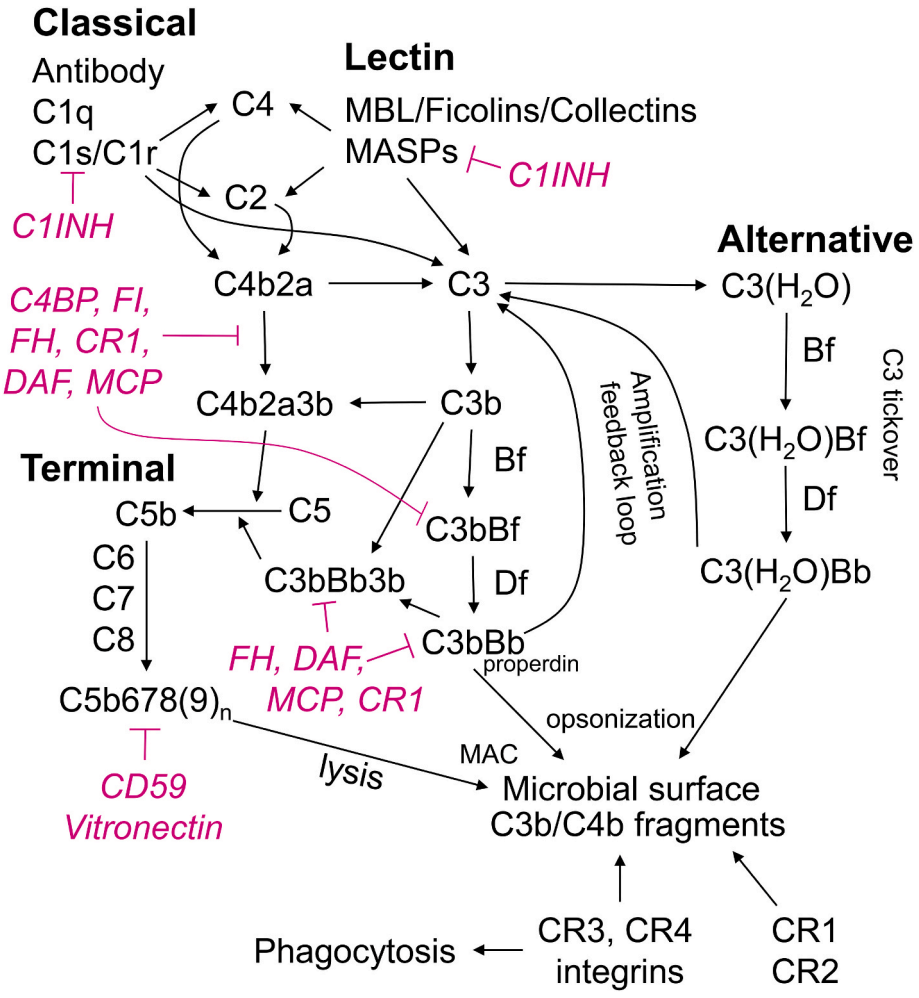



**Fig. 1.** The complement cascade in mammals. The diagram shows the classical, lectin, and alternative activation pathways, and the terminal pathway leading to the membrane attack complex. The regulatory proteins are indicated in magenta italics. Abbreviations are defined in Table 1.

**Table 1**
Definitions of abbreviations for domains and proteins.

| Domain or protein | Definition |
|---|---|
| A2M | α 2 macroglobulin domain |
| A2M_N_2 | *N*-terminal end of the α chain of the A2M domain |
| A2M_recep | receptor binding domain of the α 2 macroglobulin family |
| ALP | apextrin-like proteins |
| ANATO | anaphylatoxin homologous domain |
| ApeC | apextrin C-terminal |
| Bf | Factor B |
| C1INH | C1 inhibitor |
| C1Q | complement C1q domain |
| C2 | protein kinase C conserved region 2 |
| C345C | C-terminal domain of complement components C3, C4 and C5 |
| C4BP | C4 binding protein |
| iC3b | inactivated or degraded C3b |
| CCP | complement control protein |
| CFH-R | complement factor H-related |
| CLECT | C-type lectin or carbohydrate-recognition domain |
| CR1 | complement receptor type 1 |
| CR2 | complement receptor type 2 |
| CR3 | complement receptor type 3 |
| CR4 | complement receptor type 4 |
| CUB | complement C1r/C1s, Uegf, Bmp1 |
| DAF | decay accelerating factor |
| Df | Factor D |
| dpf | days post fertilization |
| ECM | extracellular matrix |
| EGF | epidermal growth factor |
| EGF_2 | epidermal growth factor_2 domain |
| EGF-CA | calcium-binding epidermal growth factor domain |
| EGF-like | epidermal growth factor-like domain, unclassified subfamily |
| FBG | domain at the C-terminal region of fibrinogen β and γ chains |
| FH | Factor H |
| FH-R | Factor H related |
| FI | Factor I |
| FIMAC | Factor I membrane attack complex |
| FTP | eel-Fucolectin - Tachylectin-4 - Pentraxin-1 domain; in lectins from the Japanese eel, frog pentraxin, horseshoe crab tachylectin, fly fw protein |
| GFP | green fluorescent protein |
| hpf | hours post fertilization |
| hpi | hours post infection |
| HYR | hyalin repeat |
| IL | interleukin |
| INB | integrin β *N*-terminal |
| Int_β cyt | cytoplasmic domain of β integrins |
| Int_β tail | integrin β tail domain |
| Int_α | Integrin α domain |
| KR | kringle domain |
| LDLa | low-density lipoprotein receptor domain class A, includes a cys–rich repeat with 6 cys |
| LU | Ly-6 antigen/uPA receptor-like |
| MAC | membrane attack complex |
| MACPF | membrane attack complex and perforin-like |
| MBL | mannose binding lectin |
| MCP | membrane cofactor protein |
| MASP | mannose binding lectin-associated serine protease |
| NK | natural killer |
| PAN_AP | divergent subfamily of APPLE domains |
| PAMP | pathogen-associated molecular pattern |
| PF | perforin-like |
| Pfam A2M_comp | complement component region of the α 2 macroglobulin family |
| Pfam A2M_N | α 2 macroglobulin domain *N*-terminal |
| Pfam Collagen | collagen helix |
| PRR | pathogen recognition receptor |
| RCA | regulators of complement activation |
| serpin | serine protease inhibitor |
| SMART | simple modular architecture research tool |
| SMC | secondary mesenchyme cell |
| SP | signal peptide |
| SR | domain with many serines and arginines |
| SRCR | scavenger receptor; cysteine rich |
| Tecps | thioester-containing proteins |
| Thiol-ester_cI | thiol-ester bond-forming region; thioester site |
| TLR | Toll-like receptor |
| TM | transmembrane |
| TSP1 | thrombospondin type 1 repeat |
| Tryp_SPc | trypsin-like serine protease domain |
| vWF-A | von Willibrand Factor type A |
| WMISH | whole mount *in situ* hybridization |
| ZP | zona pellucida domain |

generate the soluble alternative pathway version of the surface bound C3 convertase ($C3(H_2O)Bb$). Although the alterative pathway can function in surveillance for pathogens, its essential function is the positive amplification loop that augments the speed of complement activation and target opsonization (Fig. 1). The essential value of these activation pathways is to coat foreign cells or molecules with C3b or C4b, of which some participate in convertase complexes while others act as opsonins to augment phagocytosis and destruction of the pathogen through cellular degradation processes.

In addition to augmenting opsonization by cleaving and generating additional C3b fragments, C3 convertase complexes can incorporate an additional C3b fragment, which broadens the protease function to include cleaving and activating C5 (Fig. 1). The larger C5b fragment initiates the terminal pathway, and C5a acts as a potent anaphylatoxin that augments inflammation. C5b associates with terminal components that share some domain organization, C6 and C7, followed by the addition of C8 that associates with the membrane of the foreign cell. This complex promotes the multimerization of C9 that inserts through the membrane and forms the membrane attack complex (MAC) (Kolb and Muller-Eberhard, 1975; Sonnen and Henneke, 2014), which is a large pore of 10 nm through which molecules pass in and out of the cell without regulation and is lethal (Fig. 1). Pore formation is facilitated by the MAC and perforin-related (Macpf) domains that are signature domains of the C6-C9 proteins. In summary, the key aspects of the mammalian complement system are i) covalent and permanent bond formation between thioester proteins and foreign surfaces, ii) opsonization of cells and molecules that tags them as foreign, iii) recognition of the opsonin tags by phagocytes leading to phagocytosis and effective killing of foreign cells, and iv) lysis of foreign cells by the MAC.

## 2. Historical background of complement components in animals

Up to about the early 1970s, most immunologists believed that only vertebrates had or needed an immune system and that all vertebrates made antibodies (reviewed in (Marchalonis, 1977)). This belief correlated with the assumption that antibodies were required to activate the complement cascade (Mayer, 1973). At the time, the alternative (or properdin) pathway had been identified (Pillemer et al., 1954) and did not require antibody activation, but its biological relevance was not understood. It was proposed as the ancestral complement system (Lachmann, 1979) because it was assumed that the vertebrate complement system required antibody activation (reviewed in (Reid and Porter, 1981)). Furthermore, the lectin pathway would remain unknown until 1990 (Lu et al., 1990; Ohta et al., 1992). To understand the origins of immunity, investigations in the late 1960s and 1970s focused on evaluating antibodies in the cyclostomes or jawless fish (lampreys and hagfish) (Acton et al., 1971; Raison et al., 1978a, b) and the chondrichthyes (sharks and rays) (Marchalonis and Edelman, 1968), which have been extended more recently (Matz et al., 2021; Pettinello and Dooley, 2014)). Because both of these groups of fish are positioned at the base of the vertebrate clade (Miyashita et al., 2019; Romer, 1967), they were the optimal species to provide evolutionary insight into the appearance of the adaptive immune system in vertebrates. It was assumed at the time that these fish had vertebrate-like antibodies that were required for complement function. The lamprey antibody was

within the same size range as that for dogfish (shark) and ray antibodies, although it had the unusual characteristic of dissociating into heavy and light chains during isolation in the absence of reducing agents (Marchalonis and Edelman, 1968). Furthermore, the light chain was not reproducibly identified on gels, which could not be explained (Raison et al., 1978a). However, a hagfish cDNA sequence predicted to encode antibody, and the *N*-terminal sequence of the assumed hagfish antibody both showed that these molecules matched to C3 (Hanley et al., 1992; Ishiguro et al., 1992). Furthermore, functional analysis of hagfish C3 demonstrated its opsonic activity to induce phagocytosis by hagfish leucocytes (Raison et al., 1994). C3 homologues were also reported for the lamprey (Nonaka et al., 1984). These results were strong inference that cyclostomes do not make vertebrate-like antibodies, challenging the paradigm that all vertebrates have antibodies as key proteins in their immune systems. It was also assumed at the time that antibody activation of the classical pathway was essential (reviewed in (Reid and Porter, 1981)) and that the discovery of a complement component in the cyclostomes and the absence of antibodies was unexpected.

In line with the notion that all vertebrates have adaptive immune systems that function with antibodies, it was also believed in the late 1970s that invertebrates did not have immune systems of any kind, although this was under debate. Initial demonstrations of functional immunity employed allograft rejection as an easily visualized illustration of non-self recognition and the effector functions required to reject the allogeneic tissue. This phenomenon was reported for multiple species in the echinoderm phylum including the leather star, *Dermasterias imbricata* (Karp and Hildemann, 1976), the horned sea star, *Protoreaster nodose*, the sea cucumber, *Cucumaria tricolor* (Hildemann and Dix, 1972), the painted sea urchin, *Lytechinus pictus* (Coffaro and Hinegardner, 1977), and the purple sea urchin, *Strongylocentrotus purpuratus* (Coffaro, 1979). Although allorejection demonstrated the presence of an immune system in echinoderms including attempts to show specific immune memory (Coffaro, 1980), subsequent analysis of allorejection of echinoderm tissues showed that the immune system functioned solely on innate immunity (Smith and Davidson, 1992). This led to the assumption that if the echinoderms do not have adaptive immune functions, they do not produce antibodies or have any attributes associated with antibodies including a complement cascade. However, in contradiction to this assumption, a series of studies on the green sea urchin, *Strongylocentrotus droebachiensis*, showed that phagocytosis of sheep red blood cells (SRBCs) by the large phagocyte class of coelomocytes can be augmented when anti-SRBC IgM plus mammalian C3 are bound to the SRBCs (Bertheussen, 1981; Bertheussen and Seljelid, 1982; Kaplan and Bertheussen, 1977). In the absence of mammalian C3 or after deactivation of mammalian Bf with heat, IgM bound to SRBCs does not augment phagocytosis by the coelomocytes (Bertheussen, 1982b; Bertheussen and Seljelid, 1982). This implies that coelomocytes have a cell surface complement receptor capable of binding mammalian C3b or iC3b to induce the uptake of the SRBCs. This, in turn, infers that echinoids express a C3 homologue and have a complement system. Furthermore, analysis in the sea star, *Asterias forbesi*, suggested a two component complement system composed of a C3 homologue and a second heat labile component with serine protease activity consistent with a Bf homologue (Leonard et al., 1990).

In 1979, Lachman proposed a primitive or archaeo-complement system that is auto-activated by C3 tick-over or cleavage by microbial proteases to initiate an alternative-like pathway that functions in opsonization (Lachmann, 1979). These predictions of an echinoderm complement system were confirmed when an expressed sequence tag (an early version of RNAseq) identified a partial sequence of a C3 homologue (Smith et al., 1996) followed with a full length cDNA sequence (Al-Sharif et al., 1998). These reports described a deduced SpC3 homologue with a single processing site to generate a two-chain protein and verified the presence of a complement gene in the purple sea urchin, *Strongylocentrotus purpuratus*, with expression in coelomocytes that could be induced by immune challenge (Clow et al., 2000). Furthermore, a Bf analogue was also reported (Smith et al., 1996) and subsequently as a full length cDNA that is expressed constitutively by coelomocytes (Terwilliger et al., 2004). SpBf has the expected domains for a Bf analogue including complement control protein (CCP) modules, a von Willibrand Factor type A (vWF-A) domain, a serine protease domain, and a conserved cleavage site for a putative factor D-like protease (Smith et al., 1998). The archaeo-complement system predicted by Lachmann was essentially correct, but was modified and expanded from direct C3 activation to a minimal model with a proposed alternative pathway. The minimal complement system included a lectin based activation system with an amplification feedback loop consistent with an alternative pathway, plus core proteins with a protease to activate Bf and augment C3 activation (Smith et al., 1999, 2001). In addition to the activation mechanisms, a complement cascade must be balanced with regulatory proteins to dissociate the C3bBb convertase and degrade C3b that terminates the complement activity. These reports clearly established the complement system in animals that survive in the absence of an adaptive immune system including antibodies to activate the classical pathway. Based on the initial report of C3 in echinoderms, subsequent investigations have identified C3 homologues in a wide range of invertebrates including a tunicate (Nonaka and Azumi, 1999), a gorgonian coral (Dishaw et al., 2005), and an anemone (Kimura et al., 2009).

## 3. Methods to explore the echinoid complement system

Recent advances in genome sequencing, annotation and the availability of protein domain databases provide the opportunity to expand our understanding of invertebrate complement systems. We employed a protein domain-based approach to search improved and newly available Echinoderm genomes (as of 2022) to identify candidate proteins and genes with matches to complement components. These results provide a hypothetical outline of the structure and putative functions of the echinoid complement system, which may be used as a foundation for future functional analyses.

### 3.1. Searches of deduced proteins encoded by genes and cDNAs identifies echinoderm complement proteins

The slow discovery of the complement components in echinoderms based on sequences of single cDNAs was accelerated significantly with the availability of the purple sea urchin genome sequence (Sodergren et al., 2006) in which searches and the initial annotation of immune genes included complement analogues (Hibino et al., 2006; Rast et al., 2006). Since then, searches of genes and proteins can be carried out through Echinobase.org (Arshinoff et al., 2022) or NCBI.gov. Because the scientific literature on complement in echinoderms is sparse, proteins were identified based on BLAST searches of deduced amino acid sequences in echinoderm databases using known complement proteins or domains from humans or a few other vertebrates, and in a few cases from other invertebrates. Protein matches were evaluated with the simple modular architecture research tool (SMART) in normal mode (http://smart.embl-heidelberg.de/), a domain analysis tool, to identify proteins with domains of the expected types, numbers, and organization consistent with complement proteins. SMART domain results were used to produce many of the figures to illustrate putative complement proteins in the purple sea urchin, in a few other echinoderm species, and show the human proteins for comparison. Most of these matches have been defined as analogous to mammalian complement proteins because functional data and phylogenetic analyses are generally unavailable and we made no attempts to evaluate function for many of the protein matches that we report. However, for a few cases, expression data and protein function are described when reported in the literature.

### 3.2. RT-PCR and sequencing to evaluate and verify complement gene expression

Total RNA was isolated from non-activated coelomocytes and the axial organ from *S. purpuratus* using the RNA NucleoSpin RNA XS kit (Macherey-Nagel, Düren, Germany). The RNA concentration and quality was determined by the $A_{260/280}$ ratio. RNA integrity was evaluated on 1.5% agarose gels with ethidium bromide (Sigma Aldrich). The first strand of cDNA was reverse transcribed from 1 μg of total RNA at 42°C for 1 h in a 20 μl reaction mixture containing 1 μl of ImProm-II Reverse Transcriptase (Promega) and 0.5 μg oligo (dT)-anchor primer or random primers (Promega). Specific primers were designed based on selected sequences and used to amplify the cDNAs (Table S1). PCR reactions were carried out in 25 μl with 100 ng of cDNA, 2.5 μl of 10x incubation buffer that included 15 mM $MgCl_2$ (PCR Biosystems), 0.25 μM of each primer, 10 mM of each dNTP, and 2 units of PCRBIO Classic Taq polymerase (PCR Biosystems). PCR was performed on a iCycler thermocycler (Bio-Rad Laboratories) with the following program: 94°C for 2 min, then 35 cycles of 94°C for 30 s, 55–60°C for 30 s, 72°C for 40 s, and 72°C for 10 min. Amplicons were separated on 1.5% agarose gel, bands were cut from the gel, purified with ULTRAPrep Agarose Gel Extraction MiniPrep kit (AHN Biotechnologie), ligated into pGEM-T Easy Vector (Promega), and transformed into One Shot™ TOP10 Chemically Competent *E. coli* (ThermoFisher). Clone inserts were sequenced from one or both ends to generate full length amplicon or transcript sequences.

## 4. The core proteins of the sea urchin complement system

### 4.1. The thioester-containing proteins

The core complement gene homologue, *SpC3*, is present in the sea urchin genome sequence and encodes a protein with mostly the same domains in the same order as human C3 (Fig. 2A and B). However, there are multiple *S. purpuratus* thioester-containing proteins (SpTecps) encoded by multi-copy genes in the genome rather than single copy as was initially reported (Al-Sharif et al., 1998), and all include the consensus sequence for a thioester site, GCGEQ. The initial deduced C3 homologue, now identified as SpC3-1, and a second copy, SpC3-2, are of similar sequence and domain structure, and are also similar to the human C3 protein (Fig. 2A–C). They are encoded in separate genes and *SpC3-1* is expressed in the adults, whereas *SpC3-2* is expressed in larvae (Al-Sharif et al., 1998; Hibino et al., 2006; Shah et al., 2003). *SpC3-1* is expressed by the phagocyte class of coelomocytes in adult sea urchins (Smith et al., 2006), and SpC3-1 production is augmented *in vivo* by incubating coelomocytes with LPS (Clow et al., 2000, 2004; Gross et al., 2000). Increased expression of the *SpC3-2* gene in embryos coincides with gastrulation and can also be augmented by LPS or exposure to the marine bacterial species, *Vibrio diazotrophicus* (Shah et al., 2003). Although the anaphylatoxin domain (ANATO), which is the C3a fragment that is cleaved from the α chain upon activation, is not identified in the sea urchin proteins by the SMART program, alignments indicate that this *N*-terminal end of the α chain is conserved based on the locations of five cysteines (Al-Sharif et al., 1998). Both SpC3 proteins have a conserved cleavage site for processing the pro-proteins into the α and β chains prior to secretion (Al-Sharif et al., 1998; Clow et al., 2000; Smith et al., 1999) suggesting that they are C3 homologues. There are no SpTecps encoded in the *S. purpuratus* genome with a second conserved cleavage site that is typical of a C4 protein, which is processed to three chains. There are several other genes that encode SpTecps, of which one is α 2 macroglobulin (SpA2M) that is of similar size as vertebrate A2M and includes all of the expected domains in the correct order (Fig. 2D, E).

There are a number of additional genes encoding SpTecps that may not be SpA2M analogues even though they may contain some or all of the expected domains (Hibino et al., 2006). Three include all of the expected domains for SpA2M analogues, however they are too large and include additional domains in the C-terminal region (XP_785018, XP_030832873, XP_799248). One of the longer proteins, SpTecp2 (Fig. 2F), is expressed in secondary mesenchyme cells (SMCs) of the *S. purpuratus* embryo, coincident with the onset of gastrulation (Fig. 2G). *SpTecp2* expression in the SMCs is maintained as these cells differentiate to immunocytes called pigment cells in *S. purpuratus* larvae, and migrate from the vegetal plate and the archenteron to the ectoderm (Fig. 2H, I) (Ho et al., 2016). *SpTecp2* expression is induced in those cells in response to larval incubation with *Vibrio dizatrophicus*. Two other deduced SpTecps (XP_011676138; XP_030845260) are short and missing some of the expected domains from the C-terminal region (not shown). All of the SpTecps in *S. purpuratus* have conserved thioester sites, inferring the absence of a C5 homologue that has a mutated thioester site in vertebrates, but the functions of these SpTecps are unknown.

The activity of the thioester site in SpC3-1 and the other SpTecps in the coelomic fluid of adult *S. purpuratus* have been evaluated using a standard assay of incubation with $^{14}$C-methylamine that binds to and labels active thioester sites. SpC3-1 plus a larger protein consistent in size with SpA2M both bind $^{14}$C-methylamine and therefore demonstrate functional thioester activity (Smith, 2002). These results also suggest that either the other SpTecps are not expressed in adults, are not secreted into coelomic fluid, or they do not have functional thioester sites. SpC3-1 opsonizes Baker's yeast, *Saccharomyces cerevisiae*, and induces augmented phagocytosis by sea urchin coelomocytes (Clow et al., 2004), demonstrating the key and essential function of the sea urchin complement system of opsonization and uptake of foreign cells by the numerous large phagocytes in the coelomic fluid.

### 4.2. The factor B/C2 proteins

#### 4.2.1. Factor B/C2 proteins in vertebrates

Bf and C2 proteins in humans are composed of three CCP modules, a vWF-A domain and a serine protease domain (Fig. 3A, I), which function through the formation of an active site in the protease domain by folding a histidine, an aspartic acid, and a serine into close proximity (Polgar, 2005). In the human genome, *Bf* and *C2* are single copy genes, are tightly linked, and their sequence similarity suggests gene duplication (Campbell et al., 1984). Although they are similar, Bf binds to C3b, which exposes a scissile bond that is cleaved by factor D (Df) (Hourcade and Mitchell, 2011) to form the alternative pathway C3 convertase. C2 functions similarly by binding C4b, but is cleaved by C1s to form the classical and lectin pathway C3 convertase (Fig. 1).

### 4.3. Factor B proteins in sea urchins

Sea urchin SpBf has the expected domain structure relative to vertebrate Bf with multiple CCP modules, a vWF-A domain, and a serine protease domain (Fig. 3B) (Smith et al., 1998). The sequence of SpBf also has a conserved factor D cleavage site as well as a $Mg^{2+}$ binding site. The major difference between human Bf and SpBf is that the latter has five CCPs, rather than the expected three. However, this variation is not uncommon among Bf proteins in other deuterostomes such as CcBf from the common carp, *Cyprinus carpio*, that has four CCPs (Fig. 3G) (Nakao et al., 2002) with an apparent duplication of CCP1 (Terwilliger et al., 2004), and AjBf from the Japanese sea cucumber, *Apostichopus japonicus*, that has five CCPs (Fig. 3H) (Zhong et al., 2012). Of the five CCPs in SpBf, CCP1 and CCP2 are most similar to vertebrate Bf CCP1, SpBf CCP3 is most similar to vertebrate CCP2, SpBf CCP4 is most similar to vertebrate CCP2 and CCP3, and SpBf CCP5 is most similar to vertebrate CCP3 (Terwilliger et al., 2004). Phylogenetic analysis places SpBf at the base of the Bf phylogenetic tree predating the Bf/C2 duplication event (Smith et al., 1998, 2001), suggesting that either five CCPs is the ancestral state or five CCPs appeared in SpBf through exon duplications after the divergence from chordates (Terwilliger et al., 2004).

*SpBf* is expressed in the large phagocyte class of coelomocytes and not in the other coelomocyte types (Smith et al., 2006). However, unlike *SpC3*, *SpBf* shows constitutive expression that is not induced by immune

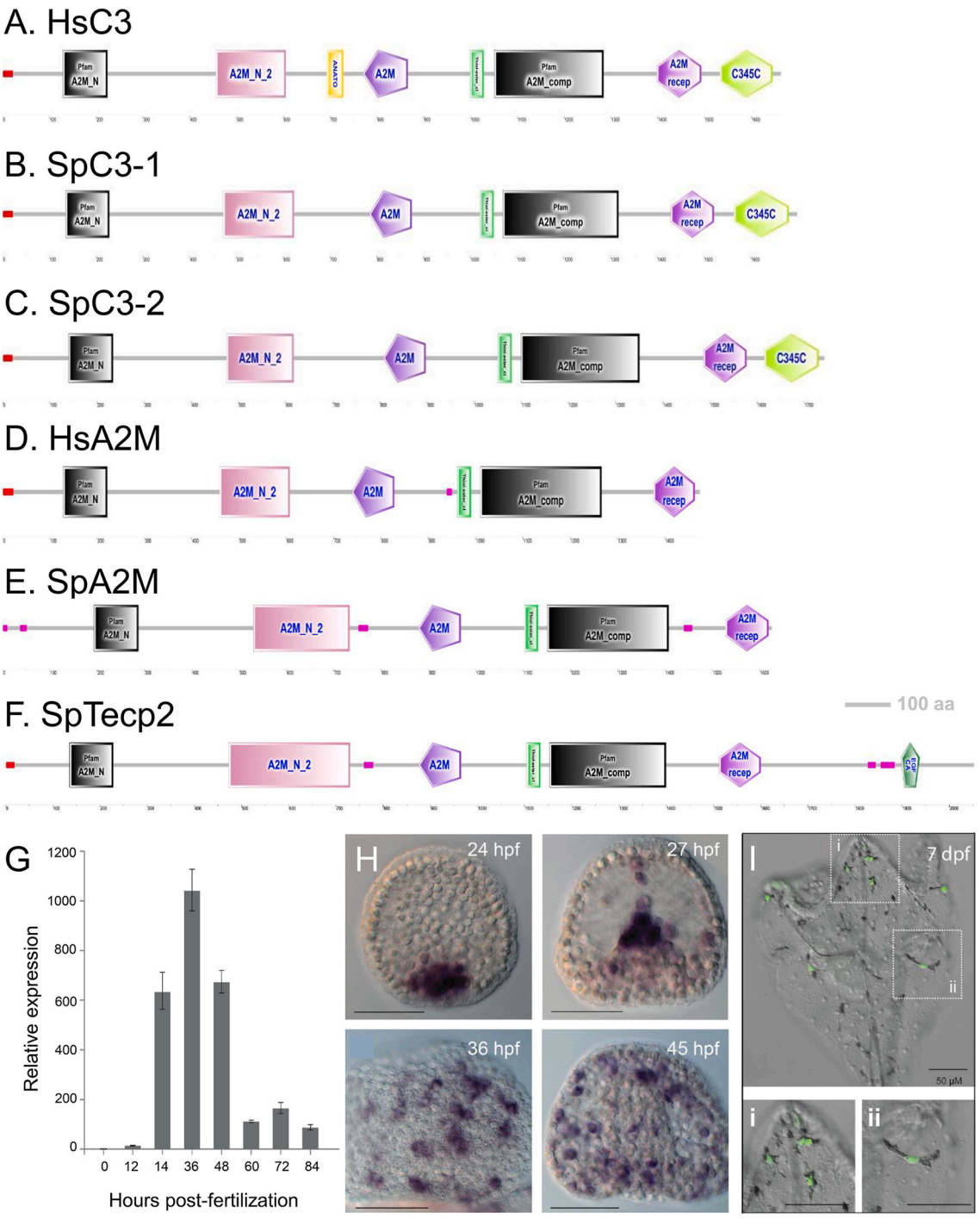



**Fig. 2.** The thioester-containing proteins in *Strongylocentrotus purpuratus*. The genome of the purple sea urchin encodes several thioester-containing proteins (SpTecps). There are two SpC3 proteins, one α 2 macroglobulin (SpA2M) protein, and several SpTecps. The domain structure of the sea urchin proteins are shown with the human proteins for comparison (see Table 1 for definitions of domain abbreviations). A. Human C3 (AAA85332.1). B. *S. purpuratus* C3-1 (AAC14396.1). C. *S. purpuratus* C3-2 (XP_780931.4). D. Human α 2 macroglobulin (EAW88590.1). E. *S. purpuratus* α 2 macroglobulin (XP_030830141). F. *S. purpuratus* Tecp2 (XP_799248.3). The additional SpTecps encoded in the sea urchin genome have domains that do not correspond to either C3 or A2M and are not shown. The truncated SpTecp1 (XP_030845260.1) has a single predicted A2M domain and a single thioester-like domain. SpTecp5 (XP_011676138.2) is similar in structure to SpTecp2 but is missing a few C-terminal domains. See main text for details. See Table 1 for definitions of domain abbreviations. G - I. *SpTecp2* is expressed in pigment cells (immunocytes) of *S. purpuratus* during embryonic development. G. Relative expression of *SpTecp2* begins at early blastula stage (14 hpf) and peaks in early gastrulae (36 hpf). H. Expression is localized to the non-skeletal mesenchyme and a subset of its derivatives, the pigment cells. I. In larvae (7 dpf), an *SpTecp2:GFP* reporter construct is expressed exclusively in some pigment cells (green) that also express echinochrome A, which appears black (insets i, ii). Scale bars indicate 50 µm. G and H are modified and re-printed from Ho et al. (2016) in accordance with the creative commons license.

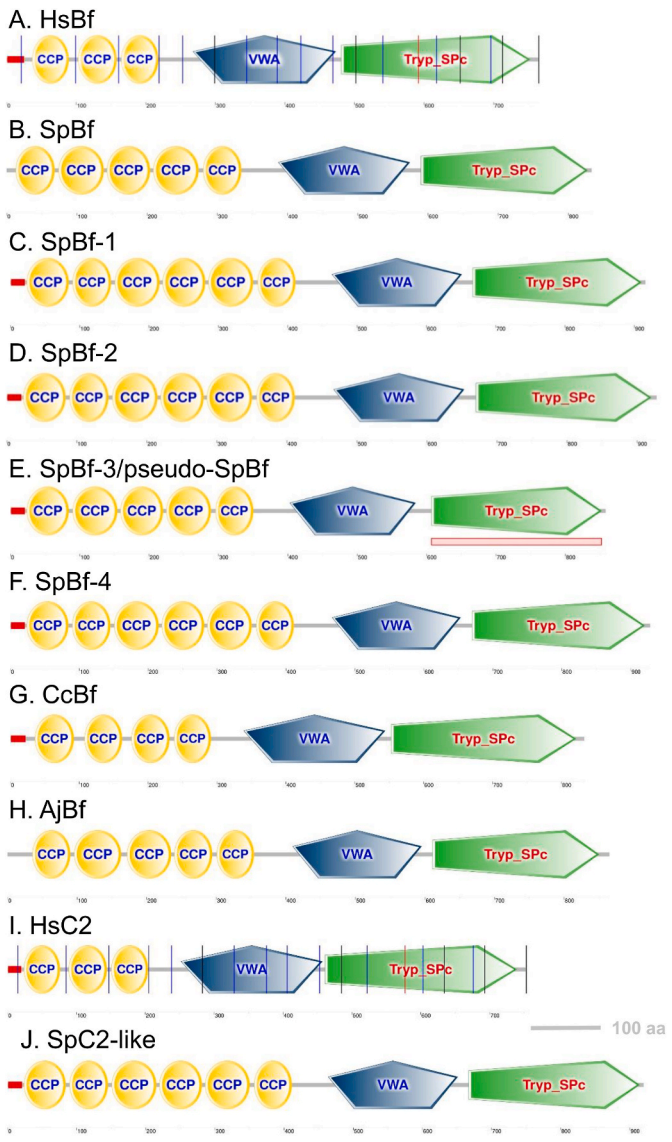



**Fig. 3.** The SpBf and SpC2-like proteins in *Strongylocentrotus purpuratus*. The domain structures of sea urchin Bf and C2 proteins are similar to those in other deuterostomes. All proteins contain CCP modules, a vWF-A domain, and a Tryp-SPc domain. Most proteins include an SP (red) at the *N*-terminus. A. Human Bf (CAA51389.1). B. *S. purpuratus* Bf (NP_999700.1). C. *S. purpuratus* Bf-1 (XP_030844905). D. *S. purpuratus* Bf-2 (XP_030844895). E. *S. purpuratus* Bf-3/pseudo-SpBf (XP_011669123). The underlined protease domain indicates a putative loss of function from a mutation that replaces the histidine with a threonine in the active site. F. *S. purpuratus* Bf-4 (XP_030853901). G. Common Carp, *Cyprinus carpio* Bf (BAA34707.1). H. Japanese sea cucumber, *Apostichopus japonicus* Bf (AEP68015.1). I. Human C2 (AQY77246.1). J. *S. purpuratus* C2-like (XP_030853760.1). Vertical lines designate exon/intron boundaries when provided in the SMART output. See Table 1 for definitions of domain abbreviations.

challenge (Terwilliger et al., 2004). This expression pattern is similar to that of vertebrates in which *Bf* expression levels are relatively constant and at a lower level than that of vertebrate *C3* (Colten, 1985; Sackstein and Colten, 1984). When evaluating *SpBf* expression in coelomocytes, messages appear in two sizes and further analyses identified alternative splicing that deleted sequence encoding CCP1, or CCP4, or both (Terwilliger et al., 2004). Although the CCP1 deletion results in a frameshift, alternative splicing of CCP4 encodes a full-length protein. This suggests that *SpBf* messages undergo alternative splicing but does not rule out the presence of as many as three similar *SpBf* genes in the *S. purpuratus* genome. At the time, alternative splicing for *SpBf* messages seemed plausible because each CCP module is encoded by a separate exon that is surrounded by canonical splice sites. The putative outcome of alternative splicing results in messages encoding SpBf proteins with five, four, or perhaps three CCP modules.

When the *S. purpuratus* genome sequence (ver 2.1) was first annotated, two additional *SpBf* genes were identified (Hibino et al., 2006), and in the *S. pupuratus* genome (ver 5.0), a total of five *SpBf-like* and one *SpC2-like* genes are present. All deduced proteins have the typical Bf protein structure with CCPs, a vWF-A domain, and a serine protease domain (Fig. 3C–F, J). Four of the deduced proteins have six CCPs (SpBf-1, SpBf-2, SpBf-4, and SpC2-like), and two have five CCPs (SpBf and SpBf-3/pseudo-SpBf). The serine protease domain in SpBf-3/pseudo-SpBf is likely non-functional because the conserved histidine that functions in the catalytic triad is replaced with a threonine (Fig. 3E). The SpC2-like protein has six CCPs rather than three as in vertebrate C2 (Fig. 3I, H). It is unlikely that SpC2-like is a C2 analogue because no gene has been identified that encodes an SpC4 protein in the sea urchin genome (see section 4.1 above), based on our assumption that SpC2 would associate with SpC4, as in vertebrates. Therefore it is likely that SpC2-like is another SpBf protein. It is noteworthy that none of the deduced SpBf proteins have less than five CCPs, which infers that the transcripts encoding the SpBf proteins with three or four CCPs (Terwilliger et al., 2004) may be the outcome of alternative splicing to remove at least CCP4, and that this variation in CCPs was not the result of expression from other *SpBf* genes. It remains unknown how the different number of CCPs impact protein function, including their association with the two SpC3 proteins, perhaps other SpTceps, and the activity of the core proteins of the sea urchin complement system.

## 5. Factor D

### 5.1. Vertebrate factor D

Factor D (Df) is required to activate the C3 convertase composed of C3bBb that forms upon activation of the alternative pathway in vertebrates (Fig. 1) (Barratt and Weitz, 2021). It is a small protein composed of a single serine proteinase domain (Fig. 4A) with highly specific function that is limited to the cleavage site in Bf when it becomes exposed upon binding to C3b (Hourcade and Mitchell, 2011).

### 5.2. Searches for factor D in the sea urchin

A functional complement activation pathway in sea urchins that expresses SpC3 and SpBf requires cleavage of SpBf to activate the serine protease domain and generate a functional C3 convertase complex with SpC3b. SpBf has a conserved cleavage site for Df (Smith et al., 1998), which implies the presence of an SpDf-like protease encoded in the genome. Although searches for SpDf during the initial annotation of the *S. purpuratus* genome (ver 2.1) did not identify SpDf (Hibino et al., 2006), our searches of the genome (ver 5.0) using the human Df protein sequence identified three possible SpDf analogues with a single trypsin-like serine protease domain, of which two included a signal sequence (Fig. 4B–D). Two of these sea urchin proteins are listed as trypsins with similarities to chymotrypsin, and the three dimensional model of SpTrypsin-1 shows similarities of the protease domain folding structure with human Df (Fig. 4E). The identification of putative SpDf proteins may provide a basic starting point for future biochemical characterization, which will be required to demonstrate that one of these proteins may function as an SpDf analogue and can cleave SpBf when bound to SpC3.

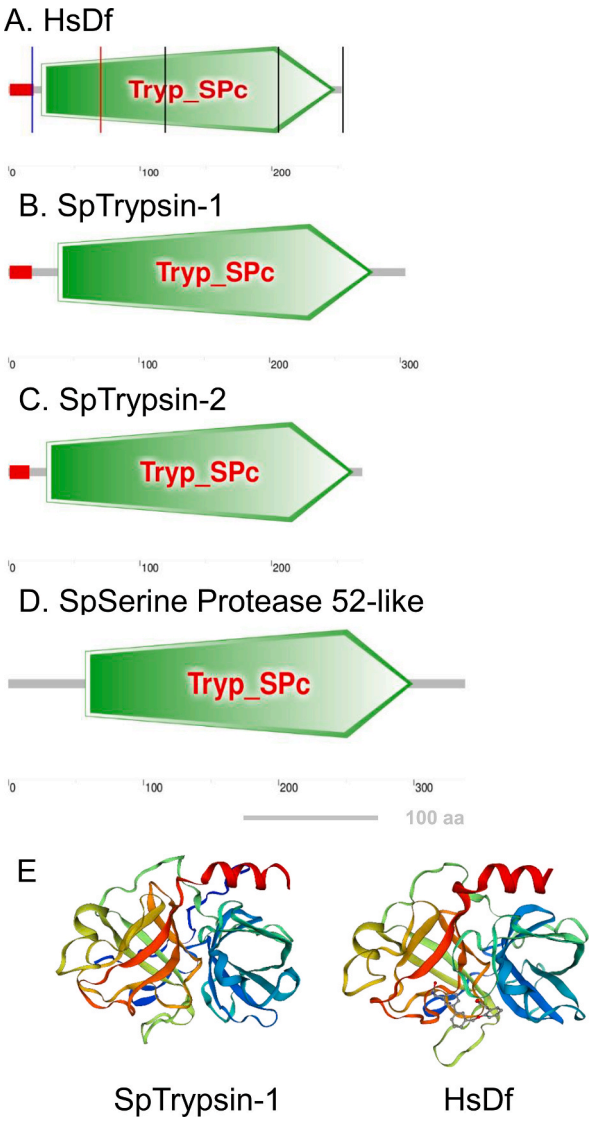



**Fig. 4.** Factor D proteins in *Strongylocentrotus purpuratus*. The purple sea urchin genome (ver. 5.0) encodes three deduced proteins that are consistent with human Df structure. The human Df protein is shown for comparison. A. Human Df (CAI2587506.1). B. *S. purpuratus* trypsin-1 (XP_001199035.1). C. *S. purpuratus* trypsin-2 (XP_030844513.1). D. *S. purpuratus* serine protease 52-like (XP_030840990.1). Vertical lines designate exon/intron boundaries when provided in the SMART output. See Table 1 for definitions of domain abbreviations. E. Estimates of three dimensional folding of *S. purpuratus* Trypsin-1 is compared to human Df. The three dimensional estimate for SpTrysin-2 is similar (not shown). Images were acquired with ColabFold (github.com/sokrypton/ColabFold).

## 6. Complement activation pathways

### 6.1. Lectin pathway activators and associated serine proteases in mammals

The activators of the lectin pathway in mammals (Fig. 1) are protein multimers composed of a collagen-like domain, a coiled coil region in some members, and a globular domain that binds foreign targets (Fig. 5A, B, D, G, I) (Hakansson and Reid, 2000). The collagen-like domain of these proteins is essential for multimerization into trimers and then into complexes of up to 18 copies forming large structures that look like a bouquet of tulips. Other proteins with similar overall structure include MBL, ficolins, and collectins (Endo et al., 2011; Hakansson and Reid, 2000; Keshi et al., 2006; Ohtani et al., 2012). MBL binds carbohydrates with terminal mannose groups that are arrayed on microorganisms in specific spacial patterns, whereas ficolins bind *N*-acetylglucosamine and do not bind mannose (Matsushita and Fujita, 2001, 2002). Both MBL and ficolins associate with MASPs resulting in activation complexes that cleave and activate other components of the lectin pathway (Fig. 1) (Matsushita and Fujita, 2001). MASP-1 cleaves and activates C3 and C2, while MASP-2 cleaves and activates C4 and C2. The three collectin proteins in humans and are identified based on their specific organ expression; liver (CL-L1, collectin 10), kidney (Cl-K1, collectin 11), and placenta (Cl-P1, collectin 12) (Hansen et al., 2016; Ohtani et al., 2012). These collectin proteins have a similar domain structure as MBL (Hakansson and Reid, 2000).

C1q also has a very similar structure to MBL and the collectins (Fig. 5A, B, D) and functions in the classical pathway in mammals because it binds the Fc portion of antibodies. However, C1q also binds directly to a variety of pathogens in the absence of antibodies to activate the lectin pathway (Bohlson et al., 2007; Reid, 2018). C1q associates with serine proteases, C1r and C1s, which are homologues of each other and the MASPs and all have similar functions (Fig. 6A) (Fujita et al., 2004b). Overall, the lectin pathway binds carbohydrates on foreign cells and activates the thioester-containing complement components that are the core and essential proteins of the complement cascade.

### 6.2. The activator proteins for the lectin pathway in echinoderms

For a lectin activation pathway to function in echinoids, proteins with the expected domain structure of a collagen-like region, an alpha helical region, and a binding head must be encoded in echinoderm genomes (Fujita et al., 2004a; Smith et al., 1999) with predicted functions of recognizing foreign structures. Furthermore, activation of the lectin pathway would require MASP association with the activating proteins. Initial searches of genes encoding lectin pathway proteins in the sea urchin genome (ver 2.1) identified some of the expected components including three SpC1q-like proteins (Fig. 5B, C, E, F) and one SpMBL (Fig. 5H) (Hibino et al., 2006). Although 46 gene models encoding fibrinogen domains were also identified, none of the deduced proteins included an *N*-terminal collagen-like domain (Hibino et al., 2006). Recent searches of the proteins encoded in the sea urchin genome (ver 5.0) verified the SpC1q and SpMBL and identified SpFicolin-1-like and SpFicolin-2-like (Fig. 5J, K). Although a collagen-like region is not identified in the SpFicolin proteins based on SMART analysis, the amino acid sequences suggest that a collagen-like region is present. These results suggest that the lectin pathway likely functions in echinoids and is involved in complement activation in these invertebrates, in agreement with previous predictions (Fujita et al., 2004a; Smith et al., 1999).

### 6.3. The associated serine proteases in echinoderms

Although the recognition components of the lectin pathway have been identified, no evidence for genes encoding proteins similar to C1r, C1s, or MASP were identified in the *S. purpuratus* genome (ver 2.1) (Hibino et al., 2006). Our searches of the genome (ver 5.0) with human C1r identified several matches to sea urchin proteins (XP_030851874.1, XP_030851875.1, XP_030851876.1; not shown), however these proteins include multiples of the expected domains in an unexpected organization and are unlikely to be analogues of C1r, C1s, or MASP. In the absence of these serine proteases in *S. purpuratus*, it is not clear how a lectin pathway may function (Al-Sharif et al., 1998; Smith et al., 1998). However, searches of all deduced echinoderm proteins available on Echinobase.org identified three C1r/C1s/MASP-like sequences in the bat star, *Pateria mineata*, and two in the crown-of-thorns sea star, *Acanthaster planci* (Fig. 6B–D). These proteins have two CUB domains, two CCP modules, one calcium-binding epidermal growth factor (EGF-CA) domain, as well as the serine protease domain. For each sea star, additional putative C1r/C1s/MASP-like proteins were identified but with unexpected domain organization. One is missing the

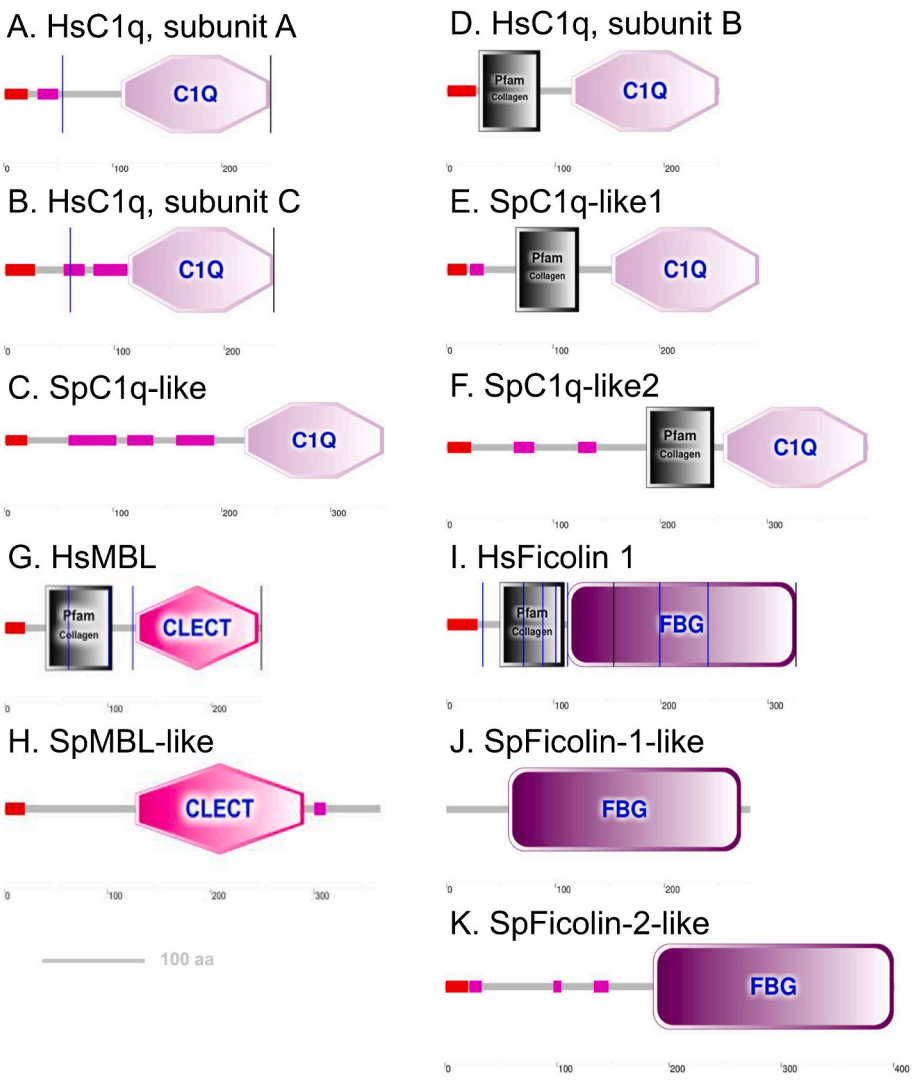



**Fig. 5.** The activation proteins of the lectin pathway in *Strongylocentrotus purpuratus*. There are four deduced proteins encoded in the genome (ver 5.0) of the purple sea urchin with domain structure consistent with activators of the lectin pathway. Three proteins have domains similar to those in C1q, one is similar to MBL, and two are similar to ficolins. The human C1q subunits, MBL, and ficolin 1 are shown for comparison. A. Human C1q subunit A (NP_001334395.1). B. Human C1q subunit C (NP_001334548.1). C. *S. purpuratus* C1q-like (XP_003729490) is annotated as collagen α (X) chain. D. Human C1q subunit B (NP_000482.3). E. *S. purpuratus* C1q-like1 (XP_003724148) is annotated as collagen α-2 (VIII) chain. F. *S. purpuratus* C1q-like2 (XP_030845267) is annotated as collagen α-1 (X) chain. G. Human mannose binding lectin C (NP_001365302.1). Human collectin 10 (CL-L1, BAA81747.1), collectin 11 (CL-K1, BAF43301.1) and collectin 12 (CL-P1, NP_569057.2) all have the same domain structure as MBL (Hakansson and Reid, 2000) (not shown). H. *S. purpuratus* mannose binding lectin-like (XP_796701) is annotated as 30 kDa spicule matrix protein. A collagen-like region is present between the SP and the CLECT domain that is not identified by SMART. I. Human Ficolin 1 (AAH20635.1 or NP_001994.2). Human Ficolin 2 (NP_004099.2) and Ficolin 3 (NP_003656.2) have the same domain structure (not shown). J. *S. purpuratus* Ficolin-1-like (XP_787024.1). K. *S. purpuratus* Ficolin-2-like (XP_781712.2). Vertical lines designate exon/intron boundaries when provided in the SMART output. See Table 1 for definitions of domain abbreviations.

*N*-terminal end including one CUB domain and the other is missing both the CUB and the EGF-CA domain (Fig. 6C, D). They are either derived from incomplete gene models that are missing the 5' end of the coding region, or they encode shorter protein versions. Based on the close evolutionary relationship of asteroids and echinoids, as well as the presence of an SpMBL and SpC1q-like proteins in echinoids, we speculate that a functional lectin pathway is present in *S. purpuratus* and that the missing C1r/C1s/MASP sequences may be based on genome annotation or assembly. Alternatively, this category of serine proteases in echinoids has different domain structures and cannot be identified with the approaches employed here. This is supported by the presence of the *C1r* genes in the two asteroid species, as well as *C1r*, *C1s*, and *MASP* genes in the ascidian *Ciona intestinalis* (Nonaka and Yoshizaki, 2004).

## 7. The terminal pathway

### 7.1. The terminal pathway in vertebrates

The terminal complement pathway in vertebrates is initiated by the C5 convertase, which is assembled and activated by one or more of the three activation pathways, and cleaves C5 into C5b and C5a. C5b associates with C6, C7, and C8 and the polymerization of C9 in the formation of the MAC, which is a large pore that results in unrestricted passage of ions and other molecules into and out of the foreign cell (Fig. 1). Foreign cells are not generally protected against MAC attack like host cells that benefit from the functions of the complement regulatory proteins (see section 9 below). The members of the terminal pathway, C6, C7, C8, and C9 are composed of a variety of domains, but the Macpf domain is the signature domain that identifies these proteins (Anderluh et al., 2014).

### 7.2. The terminal pathway in echinoderms

When the human C6, C7, C8 (including α, β and γ chains), and C9 proteins (Fig. 7B–E) are used to search the echinoderm proteins encoded in the genome (ver 5.0) on Echinobase.org or the *S. purpuratus* sequences on NCBI.gov, no convincing matches are identified. Furthermore, when C6-like proteins from either the tunicate, *Ciona intestinalis* (XP_002130807.1), or amphioxus, *Branchiostoma belcheri* (Fig. 7A), are used to search the echinoderm proteins, no matches are found with the expected domains in the expected organization as that in the invertebrate proteins or in mammalian C6. Results only identify proteins with Macpf domains (see section 7.3 below). Because a C5-like analogue is not identified (see section 4.1 above), these results infer the absence of a terminal pathway in echinoderms. This suggests that the complement system functions as an essential and efficient opsonization system that drives phagocytosis and clearance of foreign cells and pathogens, but does not have a MAC-like mechanism for pathogen lysis. However, the expansion of the Macpf-containing proteins encoded in the *S. purpuratus* genome suggests other means for cell killing.

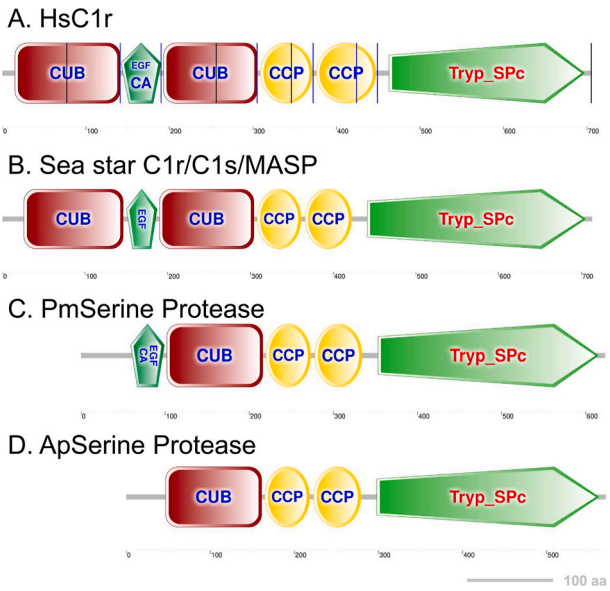



**Fig. 6.** The C1r, C1s, and MASP serine proteases in echinoderms. Searches of the deduced proteins encoded in the *S. purpuratus* genome (ver 5.0) did not result in proteins with similar domain structure as C1s, C1r, or MASP. The bat star, *Pateria mineata,* has three C1r/C1s/MASP-like proteins, and the crown-of-thorns sea star, *Acanthaster planci*, has two. Serine protease variants from the sea stars are missing the *N*-terminal region of the proteins. The domain structure for human C1r is shown for comparison, which is the same as C1s (NP_001725.1) and the MASP proteins (MASP2, NP_006601.2) (not shown). A. Human C1r (AAA51851.1). B. Several sea star serine proteases have similar structures and domains consistent with C1r/C1s/MASP including proteins from the bat star, *P. miniata* (XP_038062000.1, XP_038062005.1, XP_038062004.1) and from the crown-of-thorns sea star, *A. planci* (XP_022109261.1, XP_022109262.1). C. *P. mineata* serine protease (XP_038062004.1). D. *A. planci* serine protease (XP_022109262.1). Vertical lines designate exon/intron boundaries when provided in the SMART output. See Table 1 for definitions of domain abbreviations.

### 7.3. The Macpf proteins in echinoderms

#### 7.3.1. The genes encoding Macpf-containing proteins are expanded in echinoderms

The Macpf domain was first characterized based on the sequence similarities among the proteins of terminal complement pathway (Fig. 7B–E) that form the MAC (McCormack et al., 2013), and the perforin-like (PF) proteins (Fig. 7F) of vertebrate cytotoxic NK cells and T lymphocytes (Lowrey et al., 1989; Shinkai et al., 1988). Many Macpf-containing protein families are present across eumetazoa and are often expanded in invertebrate lineages. However, the functions of these proteins and the mechanisms of action of these repertoires remain unclear (Anderluh et al., 2014; Hibino et al., 2006; Huang et al., 2008). The initial annotation of the sea urchin genome (ver 2.1) identified 21 Macpf gene models encoding a signal peptide (SP) and a single Macpf domain (Hibino et al., 2006; Rast et al., 2006), and are organized into seven subfamilies, A - G (Fig. 7G–S). To characterize additional or divergent Macpf-containing proteins, the deduced proteins from the sea urchin genome (ver 5.0) was searched using sea urchin and human Macpf domain sequences. Updated versions of the original annotations clearly identify 19 gene models from six subfamilies, A - F, whereas the lone G family gene model no longer exists in the genome. An additional unannotated gene model (LOC756464) in the sea urchin genome (ver 5.0) encodes a protein with domain architecture most similar to members of the SpMacpfC subfamily (Fig. 7K, L). Every predicted *SpMacpf* gene family member encodes a protein with one Macpf domain, and most have an SP at the *N*-terminus. In most proteins (excluding SpMacpfA2, -A3, -A4, -D1 and -D4) the Macpf domain is paired with an Apextrin C-terminal (ApeC) domain (Fig. 7G, I–L, P–S). The ApeC domain is relatively newly described and is encoded in genes with functions related to immunity, the extracellular matrix (ECM), and protein-protein and protein-carbohydrate interactions (Li et al., 2021a, 2021b). It was first characterized in the sea urchin apextrin gene (see section 7.3.2 below), and more recently described in a diverse family of amphioxus proteins that share significant sequence similarity to the ApeC domain of the sea urchin apextrin protein (Haag et al., 1999; Huang et al., 2014). Several genes within the SpMacpfD subfamily also encode a coiled-coil domain (Fig. 7N–P).

Because of the of dual life history of the purple sea urchin that includes a feeding pelagic larva that metamorphoses into a benthic juvenile with adult morphology, the expression and function of the SpMacpf protein subfamilies may be partitioned across embryonic, larval, and adult stages. Differential subfamily expression in larval stages vs. adult tissues and coelomocytes in *S. purpuratus* has been demonstrated for other immune-related multigene families, such as the two *SpC3* genes (Gross et al., 2000; Shah et al., 2003) (see section 4.1 above), the Toll-like receptors (TLRs), the scavenger receptors (SRCRs), and the IL-17 cytokines (Buckley et al., 2017; Buckley and Rast, 2012). Such variable expression likely exists because the life stages before and after metamorphosis occur in markedly different environments and are therefore under different selective pressures.

#### 7.3.2. Expression of Macpf-containing proteins in sea urchin adults and embryos

The expression of the Macpf-containing proteins in echinoderms is known based on only two reports. The first is apextrin from the Australian purple sea urchin, *Heliocidaris erythrogramma* (Haag et al., 1999), which is likely a single analogue of the *S. purpuratus* Macpf proteins and contains the tandem Macpf and ApeC domains (Hibino et al., 2006). The *HeET-1* gene (AF049334) in *H. erythrogramma* encodes a candidate secreted ECM protein that was designated as apextrin due its novel association with apical ECM in the embryonic ectoderm (Haag and Raff, 1998; Haag et al., 1999). Apextrin is significantly upregulated in adult coelomocytes from the Australian purple sea urchin, *Heliocidaris erythrogramma*, upon challenge with bacteria suggesting an involvement in immune responses (Dheilly et al., 2011, 2012).

The second gene to be characterized for expression patterns, *SpMacpfA2,* encodes a protein with an SP and a Macpf domain (Fig. 7H). Expression of this gene occurs during embryonic development and in larval immune responses (Ho et al., 2016). *SpMacpfA2* expression initiates at mid gastrulation (Fig. 8A) where it is present in SMCs once they begin to delaminate and migrate from the end of the developing gut tube at about 48 h post fertilization (hpf) (Fig. 8B–J). The number of cells and proportion of larvae in a population that express *SpMacpfA2* increases with growth and maturation (Fig. 8B, F–I) until 10 days post fertilization (dpf) (Fig. 8J), at which point larvae have an average of 10.3 *SpMacpfA2*$^{+}$ cells. Expression of the *Macpf:GFP* reporter construct (Fig. 8K, L) demonstrates that *SpMacpfA2* is expressed in a subset of larval immune cells in the blastocoel, which are analogous to adult coelomocytes in *S. purpuratus*. *SpMacpfA2* is expressed in the mobile globular cells, which cluster at the tips of the larval arms and at the apical region of the larval body (Fig. 8M–P) (Ho et al., 2016). When sea urchin larvae are incubated with the marine bacterial pathogen, *Vibrio diazotrophicus*, that are likely ingested, there is a clearly observable gut inflammation response (Buckley et al., 2017; Ho et al., 2016). Although both *SpMacpfA2* expression and the numbers of *SpMacpfA2*$^{+}$ globular cells decrease during bacterial exposure (Fig. 9) (Ho et al., 2016), the subsequent upregulation of *SpMacpfA2* and eventual decrease in *SpMacpfA2*$^{+}$ cells both suggest the possibility that the protein is secreted directly as an extracellular effector. It remains unclear whether SpMacpfA2 and the other proteins encoded by the *SpMacpf* genes act as a bona fide part of the echinoderm complement system, however, the expression pattern and timing of *SpMacpfA2* in the immune cells in larvae is strong inference that the encoded protein has immune

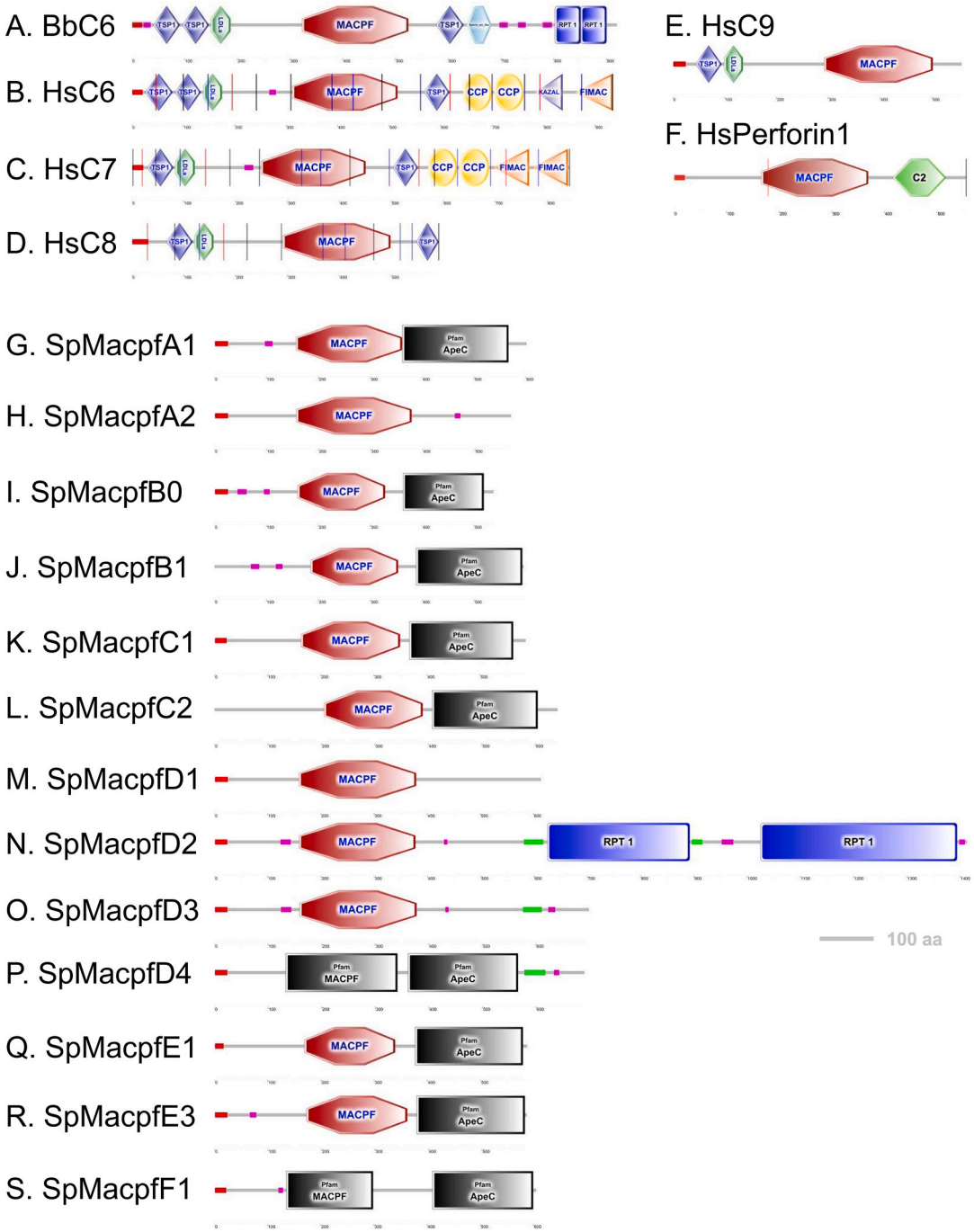
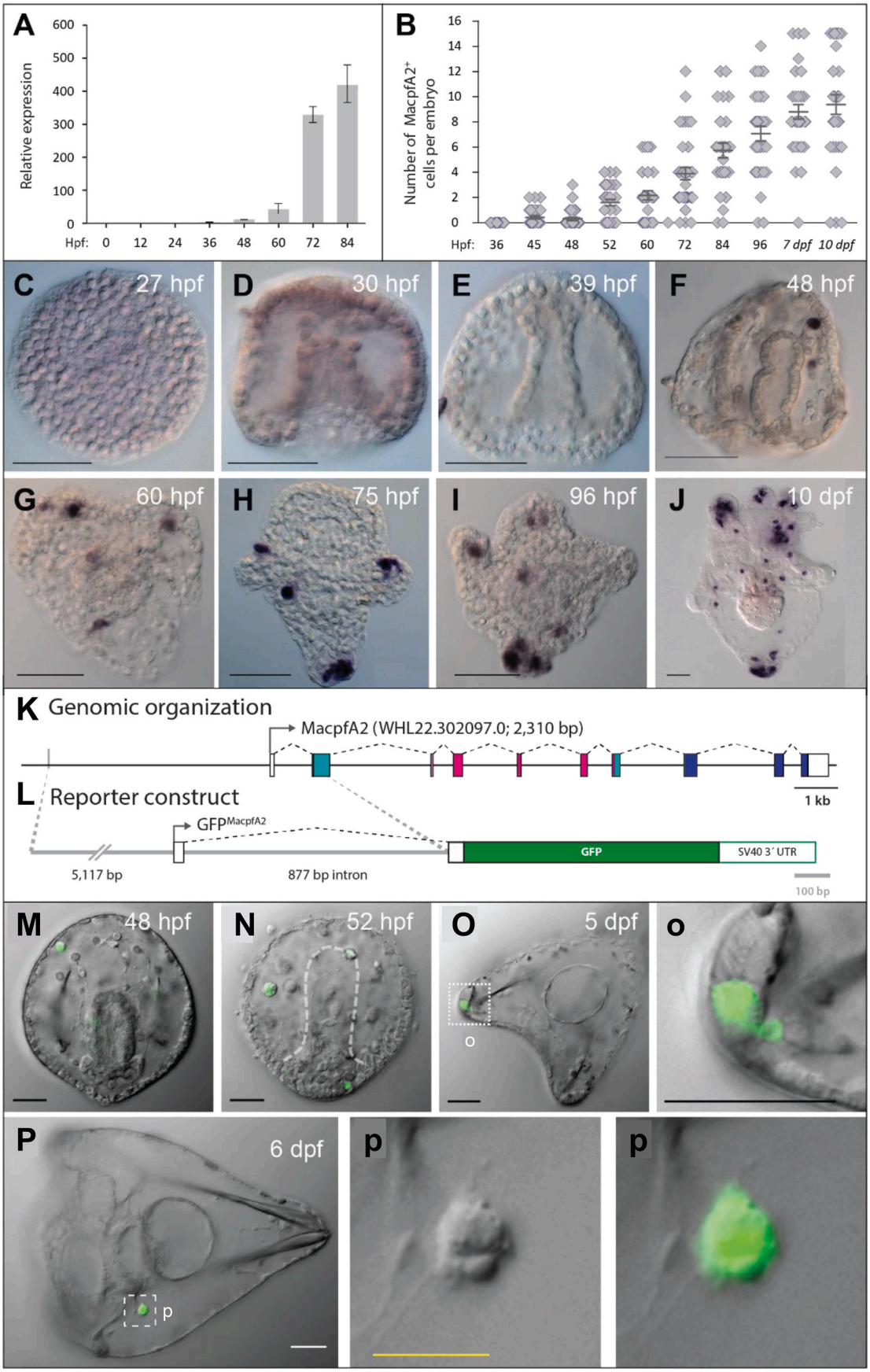



**Fig. 7.** The SpMacpf proteins in *Strongylocentrotus purpuratus*. Iterative BLAST searches using human Macpf-containing proteins or Macpf domains identify candidate *Macpf* genes in amphioxus and the sea urchin. Macpf proteins in amphioxus, *Branchiostoma belcheri*, and humans are shown for comparisons to the *S. purpuratus* proteins. A - E. Chordate complement proteins C6-C9. A. *B. belcheri* C6-like (BAB47147.1). B. Human C6 (NP_001108603.2). C. Human C7 (NP_000578.2). D. Human C8 (NP_000057.3). E. Human C9 (AAA51889.1). F. Human perforin-1 (AAA60065.1). G - H. Representative SpMacpfA subfamily domain architecture. G. *S. purpuratus* MacpfA1 (XP_011663960.2). H. *S. purpuratus* MacpfA2 (XP_030828941.1). *S. purpuratus* MacpfA3 (XP_011663960.2) and *S. purpuratus* MacpfA4 (XP_030845283.1) have the same domain structure (not shown). I - J. Representative SpMacpfB subfamily domain architecture. I. *S. purpuratus* MacpfB0 (XP_030842876.1). *S. purpuratus* MacpfB2 (XP_030842873.1) and *S. purpuratus* MacpfB3 (XP_030842872.1) have similar domain structure (not shown). J. *S. purpuratus* MacpfB1 (XP_030842877.1). K - L. Representative SpMacpfC subfamily domain architecture. K. *S. purpuratus* MacpfC1 (XP_011676478.1). *S. purpuratus* MacpfC3 (XP_030850321.1) has the same structure (not shown). L. *S. purpuratus* MacpfC2 (XP_030843828.1). An additional unannotated gene at LOC756464 in the *S. purpuratus* genome encodes a Macpf-containing protein with similar domain organization as *S. purpuratus* MacpfC2 (not shown). M - P. SpMacpfD subfamily. This is the only subfamily with predicted coiled-coil domains (green) and internal repeats (RPT1). M. *S. purpuratus* MacpfD1 (XP_784671.2). N. *S. purpuratus* MacpfD2 (XP_030842930.1) is also annotated as trichohyalin isoform X1. O. *S. purpuratus* MacpfD3 (XP_030835264.1). P. *S. purpuratus* MacpfD4 (XP_030843296.1). Q - R. Representative SpMacpfE subfamily domain architecture. Q. *S. purpuratus* MacpfE1 (XP_030837414.1). R. *S. purpuratus* MacpfE3 (XP_030837815.1). S. *S. purpuratus* MacpfE4 (XP_030837355.1) has the same structure (not shown). A previously annotated *S. purpuratus* MacpfE2 no longer exists in the genome (ver 5.0). S. *S. purpuratus* MacpfF1 (XP_030837289.1). Red at C-terminus indicates the SP, magenta indicates low complexity regions, green indicates coil-coil domains. Vertical lines designate exon/intron boundaries when provided in the SMART output. See Table 1 for definitions of domain abbreviations.




*(caption on next page)*

**Fig. 8.** The *SpMacpfA2* gene is expressed in larval immunocytes. A. Expression of *SpMacpfA2* is induced at 48 hpf and increases into the early pluteus stage (84 hpf) based on qPCR measurements of transcript prevalence. B. The number of cells expressing *SpMacpfA2* increases during development. Cells expressing *SpMacpfA2* during development were enumerated by whole mount *in situ* hybridization (WMISH) at each time point illustrating the number of *SpMacpfA2*$^+$ cells in individual embryos or larvae. Mean ± SEM are indicated. C - J. Detailed time course analysis of *SpMacpfA2* expression in cells. C - E. *SpMacpfA2*$^+$ cells are not apparent in embryos prior to 48 hpf based on WMISH results. F. At 48 hpf, one to two cells expressing *SpMacpfA2* are primarily localized near the tip of the archenteron. G - J. From 60 hpf to larval stages (10 dpf), *SpMacpfA2* expression is localized to cells within the blastocoel as well as in the larval arms and at the apex of the larval body. K - L. The *SpMacpfA2* gene organization and the green fluorescent reporter construct for *SpMacpfA2*. K. *SpMacpfA2* genomic organization. *SpMacpfA2* has ten exons, the first of which is untranslated. Untranslated regions are shown as white boxes. Magenta exons designate the coding sequence of the Macpf domain and blue exons designate the coding sequence of the ApeC domain (see Fig. 7H for SpMacpfA2 protein structure). L. The *SpMacpfA2:GFP* reporter construct includes 5.12 kb upstream of the start of transcription, plus the first intron. The GFP sequence replaces the SpMacpfA2 start of translation in exon 2. The sequences in K and L are drawn to scale (scale bars are indicated below each sequence). The non-coding region in K is expanded in L and shows the position of the inserted GFP sequence in the reporter construct. M - p. *SpMacpfA2:GFP* is expressed in globular cells. M, N. GFP$^+$ cells (green) are evident at 48 and 52 hpf. O - p. After the globular cell population differentiates, GFP is restricted to those cells in the arms, the body apex, and the blastocoel. Panel o is the arm bud region that is boxed in O and shows the GFP$^+$ globular cell. The two p panels show the globular cell outlined in panel P with the GFP overlay added in the right panel p. This figure is modified and re-printed from (Ho et al., 2016) according to the creative commons license, and from (Schrankel, 2017) as provided by the author.

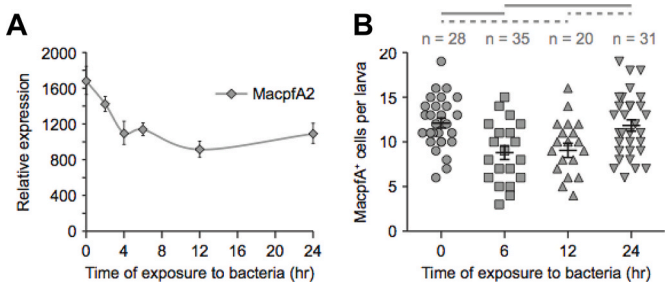



**Fig. 9.** The *SpMacpfA2* expression response varies over time during larval exposure to the bacterial pathogen, *Vibrio diazotrophicus*. A. *SpMacpfA2* gene expression decreases in the first 4 hpi when larvae are incubated with *V. diazotrophicus* (mean ± SD, n = 4 qPCR replicates). B. The number of *SpMacpfA2*$^+$ cells, as assayed by WMISH, decreases significantly during the infection time course, but eventually recovers to pre-infection numbers by 24 hpi (Student's *t*-test; horizontal lines, $p < 0.005$; dashed horizontal lines, $p < 0.05$). n, larvae counted in each sample. This figure is re-printed from Ho et al. (2016) in accordance with the creative commons license, and from (Schrankel, 2017) as provided by the author.

functions.

## 8. Complement receptors

### 8.1. Complement receptors in mammals

Mammalian complement receptor type 1 (CR1, CD35) and type 2 (CR2, CD21) are mounted on the surface of cells with the extracellular region composed exclusively of multiple CCP modules (Fearon, 1980; Gilbert et al., 2006; Weis et al., 1988). Human CR1 has 30 CCP modules (Fig. 10A) and is involved in clearing soluble immune complexes coated with C3b and C4b from circulation (Merle et al., 2021). It also has regulatory function and dissociates the C3/C5 convertase complexes and acts as a cofactor for factor I (FI) degradation of the thioester-containing proteins (Fig. 1) (Ross et al., 1982). CR2 is similar to CR1 but has 14 to 16 CCP modules, depending on the isoform (Fig. 10B) (Gilbert et al., 2006). It binds iC3b fragments and is present on the surface of B lymphocytes and functions as part of the B cell receptor for antigen. CR2 augments B cell responses and B cell receptor signaling through the detection of C3b or iC3b fragments on pathogens (Carter et al., 1988; Weis et al., 1988).

Mammalian CR3 and CR4 are heterodimers and members of the integrin protein family (Merle et al., 2021; Vorup-Jensen and Jensen, 2018). They share the β2 chain (CD18) and have different α chains; CR3 uses CD11b and CR4 uses CD11c (Merle et al., 2015a). The β chain structure has an *N*-terminal integrin β subunit (INB) domain, several EGF domains, an integrin β tail (Int_β tail) domain, a transmembrane (TM) region, and an integrin β cytoplasmic (Int_β cyt) domain (Fig. 11A). The β chains have one or more conserved NPXY motifs in the Int_β cyt domain that link to the actin cytoskeleton through binding with talin, which is essential for initiating the cytoskeletal changes required for phagocytosis. The α chain for CR3 is αM and is composed of several integrin α (Int_α) domains, a vWF-A domain, and a TM region (Fig. 11F). Although the vWF-A domain is typical of integrins that bind collagen (Whittaker et al., 2006), human CR3 binds iC3b fragments on opsonized targets (Merle et al., 2015a). The α chain for CR4 is αX that is composed of multiple Int_α domains in a β propeller repeat conformation (Fig. 11G) and binds a different iC3b fragment than CR3. Both CR3 and CR4 are involved in binding opsonized target cells and particles, which activates signaling that alters the cytoskeleton and cell shape through interactions with the β2 chain that is required for phagocytosis (Dupuy and Caron, 2008; Groves et al., 2008).

### 8.2. CR1 and CR2 receptors in echinoids

The identification of CR1 and CR2 in echinoids is addressed below with the complement regulatory proteins (see section 9.2).

### 8.3. The integrin receptors CR3 and CR4 in echinoids

Initial searches for genes encoding integrin proteins in the *S. purpuratus* genome (ver 2.1) verified three integrin β chains, SpβC, SpβG, SpβL (Fig. 11B, C) that had been reported previously (Marsden and Burke, 1997; Murray et al., 2000), and identified a fourth, SpβD (Whittaker et al., 2006). All are expressed in embryos in different cells at different times during development and SpβC is also expressed in the polygonal phagocyte type of coelomocytes (Fig. 11E). Searches of the deduced integrins encoded in the *S. purpuratus* genome (ver 5.0) identified several β integrins with domain structure similar to human β2 (Fig. 11B, C). However, the sea urchin β integrins include EGF-like domains that are not present in the human β2 integrin sequences. Each of the Spβ integrins have two NPXY motifs in the Int_β cyt domain that are consistent with binding talin and changes in the actin cytoskeleton that are required for phagocytosis (Whittaker et al., 2006). SpβC has a domain structure that is different from both human and other sea urchin β chains (Fig. 11D). It has a vWF-A domain rather than INB domain and appears to be a combination of domains that are present in human β chains and in human αM.

Searches of the deduced proteins on Echinobse.org identified α integrin proteins in *S. purpuratus*, *Lytechinus variegatus*, and *Pateria miniata* (XP_038044656; not shown) with four to six Int_α domains (Fig. 11H–J), similar but not directly comparable to the human αX protein that has three Int_α domains (Fig. 11G). None of the matches to α chains in echinoids include a vWF-A domain that is present in human αM (Fig. 11F). Results of searches indicate that the domain structures for the integrin proteins encoded in the *S. purpuratus* genome show two types of β chains (Fig. 11B–D) and two α chains with different numbers of Int_α domains (Fig. 11H, I). This leads to the hypothesis that the integrin-type


**Fig. 10.** The complement regulatory proteins with CCP modules in *Strongylocentrotus purpuratus*. The domain organization of deduced complement regulatory proteins from *S. purpuratus* are compared to the human proteins. A. Human CR1 (AAB60695.1). B. Human CR2 (NP_001006659.1). C. Human DAF (CD55) (NF06964.1). D. Human MCP (CD46) (ABK81638.1). E. Human Factor H (CAA68704.1). F. Human C4b-binding protein (NP_000706.1). G. *S. purpuratus* CR1-like1 (XP_787776.4). H. *S. purpuratus* FH-like1 (XP_030835056). I. *S. purpuratus* FH-like2 (XP_789428.3). J. *S. purpuratus* FH-like3 (XP_030845742.1). K. *S. purpuratus* FH-like4 (XP_030846616.1). L. *S. purpuratus* FH-like5L (Long variant) (XP_030847255). M. *S. purpuratus* FH-like5S (Short variant) (JT112571.1). The SP at the *N*-terminus is indicated in red and low complexity regions are indicated in pink. See Table 1 for definitions of protein and domain abbreviations.

complement receptors for the sea urchin may have different β chains, which is different from CR3 and CR4 in humans that share a β chain. We speculate that the combination of two β chains and two or more α chains may result in at least four integrin-type complement receptors. The major phagocytic cell in sea urchins is the polygonal phagocyte that is very efficient at phagocytosing foreign cells (Bertheussen, 1982a; Chou et al., 2018). Because this cell type expresses SpβC (Fig. 11E) it is tempting to speculate that, when associated with an α chain, it recognizes particles and foreign cells opsonized with SpC3b and induces phagocytosis (Clow et al., 2004).

## 9. Regulators of complement activation

### 9.1. The complement regulatory proteins in vertebrates

The thioester sites are exposed when C3 and C4 are cleaved and activated to C3b and C4b, or when C3 reacts with water to form C3 ($H_20$). The exposed thioester sites are dangerous and potentially lethal because they can form covalent bonds with hydroxyl or amine groups on any molecule including those on self cells, and can lead to cell lysis. Consequently, complement regulatory proteins in mammals are present on both cell surfaces and in extracellular fluids to protect the host from auto-attack by the complement cascade (Merle et al., 2015a). All proteins encoded within the gene cluster of regulators of complement activation (RCA) in mammalian genomes encode proteins with remarkable similarity in structure and overall composition. They consist predominantly of domains referred to as short consensus repeats (SCRs) or CCP modules that have a highly conserved globular 3D structure (Krych-Goldberg and Atkinson, 2001). Commonly, distinct functions such as inhibitory activities, binding to cell surfaces, or to other complement components can be attributed to different CCP modules in each regulatory protein. This modular organization of genes encoding repeats of CCP modules may underpin the appearance of genes encoding hybrid proteins with different arrays of CCP modules that have altered functions (Pouw et al., 2015). Several CCP-containing complement regulatory proteins with similar structures are present on cell surfaces and include CR1 (CD35), CR2 (CD21), membrane cofactor protein (MCP, CD46) and decay accelerating factor (DAF, CD46) (Fig. 10A–D) (Kim and Song, 2006). CR1, MCP, and DAF function in association with other proteins to dissociate the C3 and C5 convertase complexes to block the production of additional C3b and C5b fragments, thereby stopping the activation pathways, the feedback loop of the alternative pathway, and the terminal pathway (Fig. 1). The set of humoral regulatory proteins with multiple CCP modules includes factor H (FH) and C4 binding protein (C4BP) (Fig. 10E and F) that also function to dissociate the C3/C5 convertases. Other humoral regulatory proteins that are not composed of CCP modules include FI and C1q inhibitor (C1INH). FI is a serine protease that degrades C3b and C4b after dissociation from the convertases and uses co-factors for function including CR1, FH, C4BP, or MCP (Fig. 1). C1INH is a small protein composed of a single serine protease inhibitor (serpin) domain (Fig. 14A) that functions to dissociate the C1r/C1s/MASP proteins from C1q, MBL, and ficolins to block the

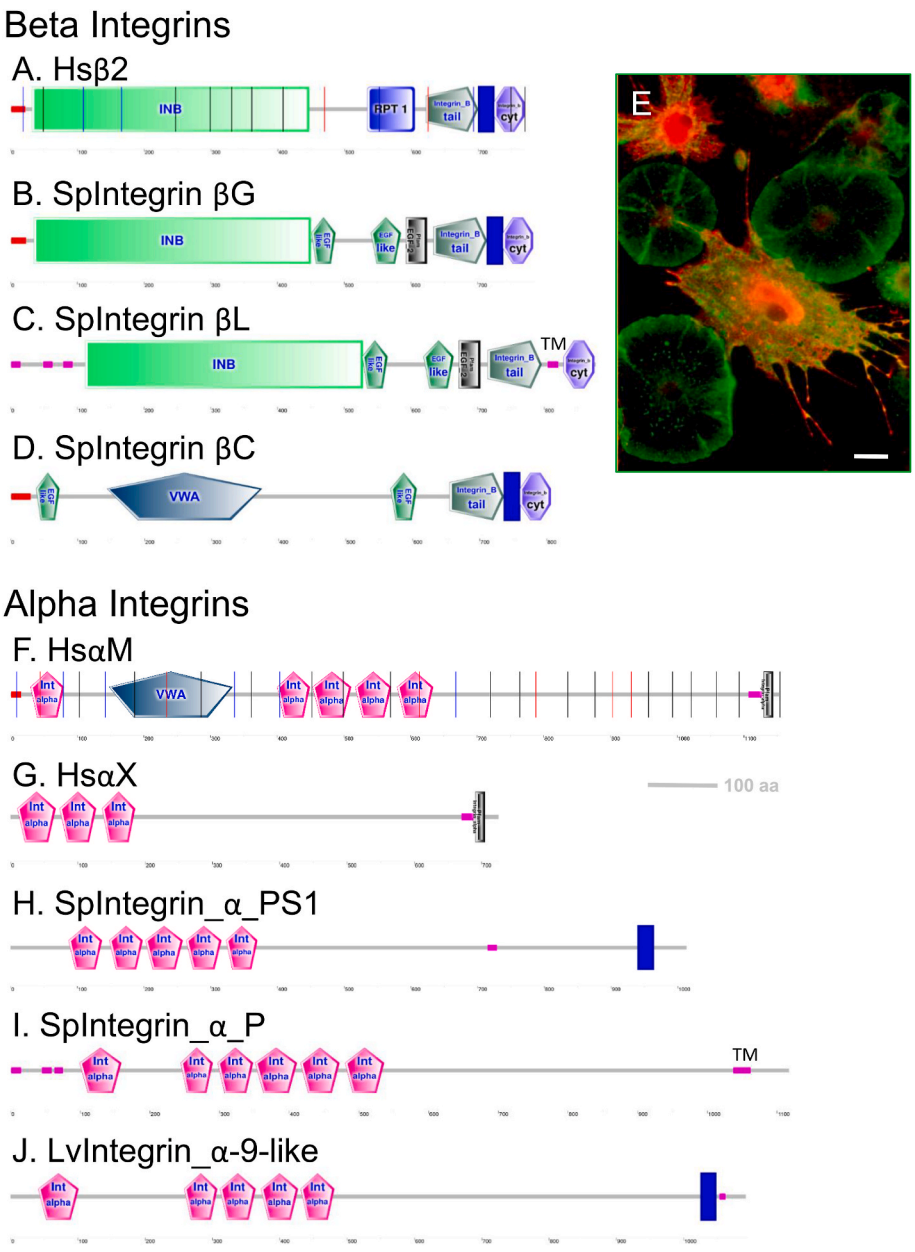



**Fig. 11.** The integrin type complement receptors in echinoids. Sea urchin genomes encode several heterodimeric integrin-like proteins with domains that are consistent with α and β chains. A - E. Beta integrins. A. Human β2 or CD18 (NP_000202.3). B. *Strongylocentrotus purpuratus* integrin βG (NP_999732). *S. purpuratus* integrin_β_1_A (XP_030845944) has the same domain structure (not shown). C. *S. purpuratus* integrin βL (NP_999731). D. *S. purpuratus* integrin βC (NP_999730 or AAB39740.2) as reported by (Whittaker et al., 2006) and (Murray et al., 2000). E. Integrin SpβC is expressed in polygonal phagocytes. A polygonal phagocyte in the center of the image expresses SpβC (orange). The discoidal cells (disc-like morphology) are also large phagocytes but do not express the SpβC integrin. The anti-SpβC antibody was provided by Robert Burke (Murray et al., 2000). The cells are counterstained for actin (green). Scale bar indicates 10 μm. The image was provided by John Henson (Dickinson College). This figure is modified and reprinted with permission by Springer/Nature/Palgrave, and appeared originally in Advances in Experimental Medicine and Biology 708:260–301. Echinoderm Immunity, LC Smith, J Ghosh, KM Buckley, LA Clow, NM Dheilly, T Haug, JH Henson, C Li, CM Lun, AJ Majeske, V Matranga, SV Nair, JP Rast, DA Raftos, M Roth, S Sacchi, CS Schrankel, K Stensvåg. © 2010. F - J. Alpha integrins. F. Human αM or CD11c (NP_001139280.1) is a subunit of CR3. G. Human αX or CD11b (EAW52140.1) is a subunit of CR4. H. *S. purpuratus* integrin_α_PS1 (XP_030846529). I. *S. purpuratus* integrin_α_P (NP_999642). J. *Lytechinus variegatus* integrin_α-9-like (XP_041477995.1). *Pateria mineata* integrin_α-9-like (XP_038044656) has the same domain structure (not shown). Integrin Spα-V (XP_030836102) from *S. purpuratus* has two Int_α domains but no TM region (not shown). TM, a putative transmembrane region, which is indicated as a low complexity region (pink), has 17 to 20 hydrophobic amino acids. Vertical lines designate exon/intron boundaries when provided in the SMART output. See Table 1 for definitions of domain abbreviations.

classical and lectin activation pathways (Fig. 1) (Bos et al., 2002). The S-protein, or vitronectin, and CD59 are cell surface regulatory proteins that function to control the terminal pathway by associating with the C5b678 complex to block the association and multimerization of C9 into the MAC and thereby avoid lysis of self cells (Fig. 1) (Kim and Song, 2006).

In addition to FH, both humans and mice express proteins that are highly similar to FH. These FH-Related (FH-R) proteins are individually encoded in the RCA gene cluster and are composed solely of multiple CCP modules like FH. Five *Complement FH-R* (*CFH-R*) genes have been reported in humans, *CFH-R1* to *CFH-R5* (reviewed by (Skerka et al., 2013)), and are arranged in tandem in the RCA gene cluster. All *CFH-R* genes encode expressed proteins in addition to two proteins from *CFH-R4*, FH-R4A and FH-R4B, based on alternative splicing (Józsi et al., 2005). Similarly, five *MmFH-R* genes in mice have been predicted *in silico*, including *MmFH-Ra* to *MmFH-Re* of which the deduced proteins, MmFH-Rb and MmFH-Rc, are expressed as alternative splice variants (Hellwage et al., 2006). *MmFH-Ra* and *MmFH-Rd* are not expressed and are assumed to be pseudogenes. The FH-R proteins lack CCP modules that are homologous to the first four CCP modules of human FH, suggesting that the FH-R proteins do not function directly for regulating complement activation. Although the functions of all mammalian FH-R proteins are not completely understood, the current hypotheses for the functions of the human FH-R proteins are that they act as de-regulators of the complement system by competing with FH for binding to cell surfaces and to C3b (Hellwage et al., 2006).

### 9.2. *The echinoderm regulatory proteins*

#### 9.2.1. *The regulatory proteins with CCP modules*

Information on the complement regulatory proteins in non-mammalian vertebrates is scarce, and is practically absent for invertebrates. Searches for and investigations of complement regulatory proteins in non-mammals and invertebrates is particularly complex because of the similarities in sequence and protein structure that leads to difficulties in predicting putative sequence similarities to those in mammals. This may be an outcome of sequence variability over evolutionary time that originated through mechanisms of gene duplication/

deletion and exon/domain shuffling, in addition to splice variants that result in genes encoding multiple CCP-containing proteins. Consequently, functional similarities among complement regulatory proteins cannot be identified based on evolutionary relatedness or sequence similarities, as is typical for other proteins. Furthermore, the CCP modules, as in SpBf proteins, may vary in number within and among species, and can include differences based on alternative splicing (see section 4.2.2). For example, in some of the sequences that we report here, we identified transcripts encoding variant numbers of CCP modules with no information as to whether these differences are due to splicing or transcription from separate genes. Consequently, efforts to identify similarities among proteins composed almost entirely of CCP modules can only be carried out by comparing the different CCP modules rather than considering the structure of the entire protein.

The several proteins with multiple CCP modules suggest evolutionary selection based on the number and organization of these modules. When invertebrate genomes or transcriptomes are searched for sequences encoding mammalian-like complement regulatory proteins, matches can include proteins encoded by alternatively spliced mRNAs, pseudogenes, and variant genes generated by exon shuffling, or errors in deducing mRNAs from genomic sequences that result in sequence kinship relationships that are difficult to reconstruct by phylogenetic analysis. Unfortunately, the Illumina platform, which is currently the most widely used sequencing technology for transcriptomes, produces short sequences that are deconvoluted to produce full length cDNAs, which is not optimal for sequences with multiple copies of the same short domain that appear as repeats. The use of long read sequencing is expected to improve the assembly of the CCP-encoding genes and transcripts and should clarify and establish the complement regulatory proteins in a wide range of animals. However, given these difficulties, we attempt to clarify the functional predictions of the CCP-containing complement regulatory proteins by identifying them from transcripts in transcriptome databases, and determining whether these mRNAs are present in two sea urchin tissues of interest based on gene expression analyses. Results focus on identifying the complement regulatory proteins in the California purple sea urchin, *Strongylocentrotus purpuratus*, and to a lesser extent in the Mediterranean purple sea urchin, *Paracentrotus lividus*. The *S. purpuratus* genome (ver 5.0) on Echinobase.org and the NCBI.gov database were searched using HsCR1 and HsFH (Fig. 10A, E) to identify echinoid proteins with repeated CCP modules including some with a TM region. Very long putative transcripts (deduced from genomic sequences) were identified encoding multiple CCP modules, but also included many other domains in random order without any similarity to known proteins and may be the result of chimeric transcripts. After all proteins with anomalous domain organization were eliminated, several deduced proteins were identified with domain organization consistent with FH, DAF, and MCP. However, for continuity with the nomenclature used for many annotated sequences with many CCP modules, all have been defined as SpFH-like even if they may have CCP modules that are similar in number to other CCP-containing proteins.

The CCP-containing proteins for *S. purpuratus* have a number of domains in common. SpFH-like1 contains an SP and five CCP modules (Fig. 10H), SpFH-like2 has an SP and six CCP modules (Fig. 10I), SpFH-like3 is missing the SP and has 12 CCP modules that are intermingled with three hyalin repeats (HYRs) (Fig. 10J), SpFH-like4 is also missing the SP and has 16 CCP modules intermingled with six HYRs, plus two putative ephrin receptor-like domains at the C-terminus (Fig. 10K). SpFH-like5L (Long variant) (Fig. 10L) has an SP and nine CCP domains whereas SpFH-like5S (Short variant) (Fig. 10M) has an SP and five CCP domains. SpCR1-like1 (Fig. 10G) has a TM region that is consistent with HsCR1, plus 11 CCP modules, however, it also has a zona pellucida (ZP) domain, which is unlike HsCR1. Comparisons among the FH-like sequences indicate that SpFH-like1 has a region of 330 amino acids near the *N*-terminal end that is similar to the same region of SpCR1-like1 (89.1% identity and 97.3% similarity; (EMBL-EBI Lalign tool; www.ebi.ac.uk/Tools/psa/lalign/). However, the first 50 amino acids at the *N*-terminus of these two proteins are different, which includes the SP in SpFH-like1 that is missing from SpCR1-like1. To determine whether these two proteins are the outcome of alternative splicing, genome analysis indicated that they are encoded by different genes positioned on different scaffolds.

Of the CCP-encoding genes, expression from 11 were selected for further investigation by PCR amplification of cDNAs derived from coelomocytes and the axial organ from *S. purpuratus* using primers for specific sequences (Table S1). Only seven are expressed in these two tissues (Fig. 12) and those with FH similarities encode SpFH-like1, SpFH-like2, SpFH-like3, SpFH-like4, SpFH-like5L, and SpFH-like5S (Fig. 10H–M). However, the expression patterns for most of these *SpFH-like* genes show differences for expression in coelomocytes compared to the axial organ. Genes encoding SpCR1-like1, SpFH-like1, and SpFH-like5L are all expressed in the axial organ, but not the coelomocytes (Fig. 12A, B), whereas the genes encoding SpFH-like3 and SpFH-like4 are expressed in both the axial organ and coelomocytes (Fig. 12A). It is noteworthy that the axial organ is reported as a site for coelomocyte hematopoiesis (Golconda et al., 2019) and that large phagocytes are present in the axial organ (Majeske et al., 2013).

When the repertoire of primers designed for *S. purpuratus* transcripts encoding CCP-containing proteins are used to amplify cDNAs from *Paracentrotus lividus* coelomocytes, only one primer pair (SpCR1FH1 For and SpCR1FH1 Rev, Table S1) results in an amplicon of about 250 bp that has significant sequence similarity with the transcripts from *S. purpuratus*. Failure of the other primer pairs to generate amplicons from *P. lividus* cDNAs may be attributed to a variety of possible reasons. Although it is probable that the CCP-containing proteins are highly variable within and among species, as suggested from results for *Ciona intestinalis* (Sommer et al., 2012), similar sequence among echinoid species indicates sequence conservation and suggests important protein function.

To ensure that some of the deduced proteins were not the outcome of alternative splicing of transcripts from the same gene, primers were designed to amplify the full-length coding region of the transcripts (Table S1). The amplicon for *SpFH-like2* appeared as a single band of the expected size (Fig. 12C) and the sequence matched that in the NCBI database. Alternatively, the amplicon for *SpFH-like5S* was shorter than expected (primer pair SpFH-like5 CDS For and SpFH-like5 CDS Rev, Table S1) and the first four of five CCP modules in SpFH-like5S were duplicated in SpFH-like5L that had a total of nine CCPs (Fig. 10L, M). To determine whether SpFH-like5L and SpFH-like5S were the outcome of alternative splicing, a range of annealing temperatures were used for this pair of primers, however, amplicons of a longer transcript were not generated (Fig. 12C). Searches of the available genomic sequences did not identify sequences to suggest the presence of a duplicated gene with fewer exons consistent with *SpFH-like5S*. Instead, searches of the Sp Transcriptome Shotgun Assembly database identified a transcript (JT112571.1) with sequence identical to that encoding SpFH-like5S suggesting that alternative splicing of transcripts from the same gene may encode both the short and long proteins. To date, the functions of different CCP modules are not known, but considering that this module is present in many of the complement regulatory proteins, in addition to Bf and C2 in vertebrates and echinoids, CCP modules in SpBf may interact with SpC3b, and CCP modules in the regulatory proteins may function to dissociate the convertases through binding competition with SpBb for SpC3b.

### 9.3. Non-CCP regulatory proteins

There are several complement regulatory proteins that function in humans and do not include CCP modules. Human factor I (FI), prior to processing into two chains, is composed of a variety of domains with an SP, a FIMAC domain, a domain with many serines and arginines (SR), two low-density lipoprotein receptor class A (LDLa) domains, and a

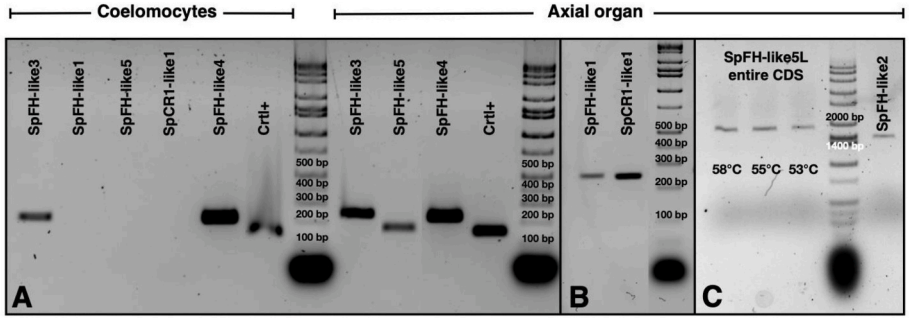



**Fig. 12.** The genes encoding putative complement regulatory proteins with CCP modules are expressed in coelomocytes and the axial organ. A, B. RT-PCR was used to verify expression of genes encoding CCP-containing proteins in coelomocytes and/or the axial organ from *Strongylocentrotus purpuratus* (see Table S1 for primer sequences). C. Different annealing temperatures are used to differentiate expression of *SpFH-like5L* and *SpFH-like5S* in the axial organ. Amplification of the entire coding region for *SpFH-like2* is also shown. Gene expression is shown for *S. purpuratus FH-like3* (JT102339.1 or XM_030989882.1), *S. purpuratus FH-like1* (XM_030979196.1), *S. purpuratus FH-like5S* or *SpFH-like5L* (XM_030991395.1), *S. purpuratus CR1-like1* (XM_782683.5), *S. purpuratus FH-like4* (XM_030990756.1), and *S. purpuratus FH-like2* (JT115129). The positive control (Crtl+) amplicon is cytoplasmic actin. Images of gels in A and B have been modified to delete blank lanes, which does not change the results that are shown. Relevant sizes of DNA standards are indicated.

trypsin-like serine protease (Tryp_SPc) domain (Goldberger et al., 1984). Protein BLAST searches for similar proteins on Echinobase.org and NCBI.gov resulted in a few echinoderm proteins with serine protease domains, but each included many other domains in different organizations that are not consistent with a vertebrate FI protein. The failure to identify an echinoderm FI is not consistent with the presence of C3 homologues because of its important functions in mammals to degrade thioester proteins that have opsonized surfaces or have been activated in solution. SpFI may have a different organization of domains and therefore cannot be identified by protein BLAST, or the gene may be present in a poorly assembled region of the genome and will become apparent as the assembly continues to be improved.

The human C1IHN protein is a small protein with a signal sequence, a low complexity region, and a single serpin domain (Fig. 13A). Searches of echinoderm proteins with the human C1IHN results in two proteins with single serpin domains that are annotated as an elastase inhibitor-like in *S. purpuratus* and a serpin B6-like in the sea star *Acanthaster planci* (Fig. 13B, C). Both echinoderm proteins are missing the *N*-terminal region relative to the human sequence, and although these have a single serpin domain similar to human C1IHN, it is only speculative that these echinoderm proteins may be C1INH analogues.

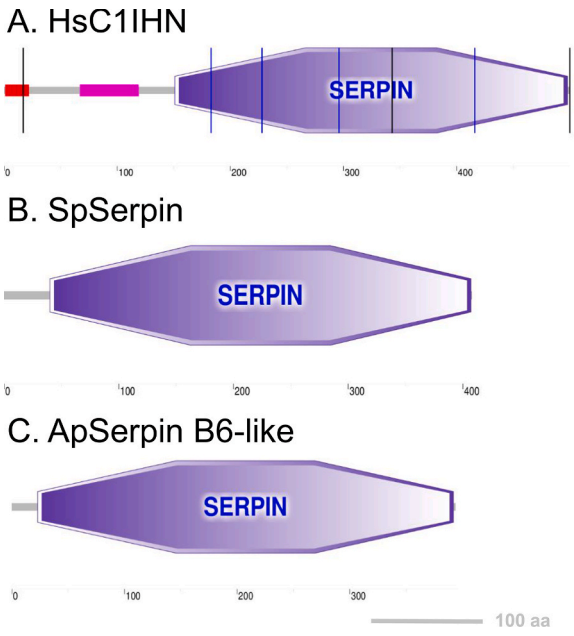



**Fig. 13.** C1INH/serpins in echinoderms. Deduced proteins with a single serpin domain that are consistent with a C1 inhibitor (C1INH) domain structure, are encoded in the genomes of *Strongylocentrotus purpuratus* and the crown-of-thorns sea star, *Acanthaster planci*. The echinoderm proteins with the most similar structure to human C1INH are shown. A. Human C1 inhibitor (NP_001027466.1). B. *S. purpuratus* serpin (XP_003730975.1) is annotated as leukocyte elastase inhibitor-like. C. *A. planci* serpin B6-like (XP_038050239.1).

Vitronectin, or S protein, is a small humoral protein with a somatomedin B domain and four hemopexin-like repeats (Preissner, 1991). It associates with the C5b67 complex to which C8 and C9 can be added, and it functions to block insertion of the complex into membranes and therefore remains in solution (Fig. 1). Vitronectin also blocks the lytic functions of perforin (Tschopp et al., 1988). Searches of echinoderm proteins encoded in the genomes at Echinobase.org with human vitronectin (NP_000629.3) identifies asteroid proteins (XP_022104779.1, XP_038046937.1) with four hemopexin-like repeats, but these proteins do not include a somatomedin B-like domain (not shown), and therefore are unlikely to be vitronectin analogues. The absence of this protein with functions to regulate the terminal pathway is consistent with the absence of genes encoding proteins in the terminal pathway in echinoids (see section 7.2).

Human CD59 regulates the terminal pathway by binding C8 that blocks binding and polymerization of C9 into the MAC (Meri et al., 1990). It is a highly glycosylated protein that associates with cell surfaes by a GPI linkage, and has a Ly-6 antigen/uPA receptor-like (LU) domain (van den Berg et al., 1995). The initial annotation of the *S. purpuratus* genome sequence (ver 2.1) reported two genes encoding SpCD59 proteins (XP_003724103, XP_030845206) (Hibino et al., 2006). However, when evaluated for domains, these proteins do not include the LU domain that is present in human CD59 (not shown). Furthermore, alignments with the human CD59 and the two reported SpCD59 sequences show little similarities except for an *N*-terminal hydrophobic region consistent with an SP, a few cysteines that align among the proteins, and a hydrophobic region at the C-terminus. Searches of echinoderm proteins encoded in the genome (ver 5.0) at Echinobase.org or at NCBI.gov do not identify matches with CD59 sequence or domain similarity. The inability to identify SpCD59 is consistent with the inability to identify vitronectin and members of the terminal pathway in the echinoderm complement system.

## 10. The theoretical complement system in echinoids

The thioester proteins are the key to understanding the functions of the complement system in echinoids. The two SpC3 homologues and the absence of SpC4 and SpC5, indicates that the classical activation pathway is absent, that the lectin pathway functions through a variety of initiating proteins, and that the terminal pathway is missing (Fig. 14). The C1r/C1s/MASP analogues are not found in the echinoid genomes, which is unexpected because they are essential for the classical and lectin activation pathways to run in the mammalian complement cascade. The echinoid MASP family proteins may have a different domain structure and therefore are not identified with the search methods employed here. However, it is noteworthy that they are present in asteroid genomes, perhaps suggesting a significant difference in the asteroid vs. echinoid

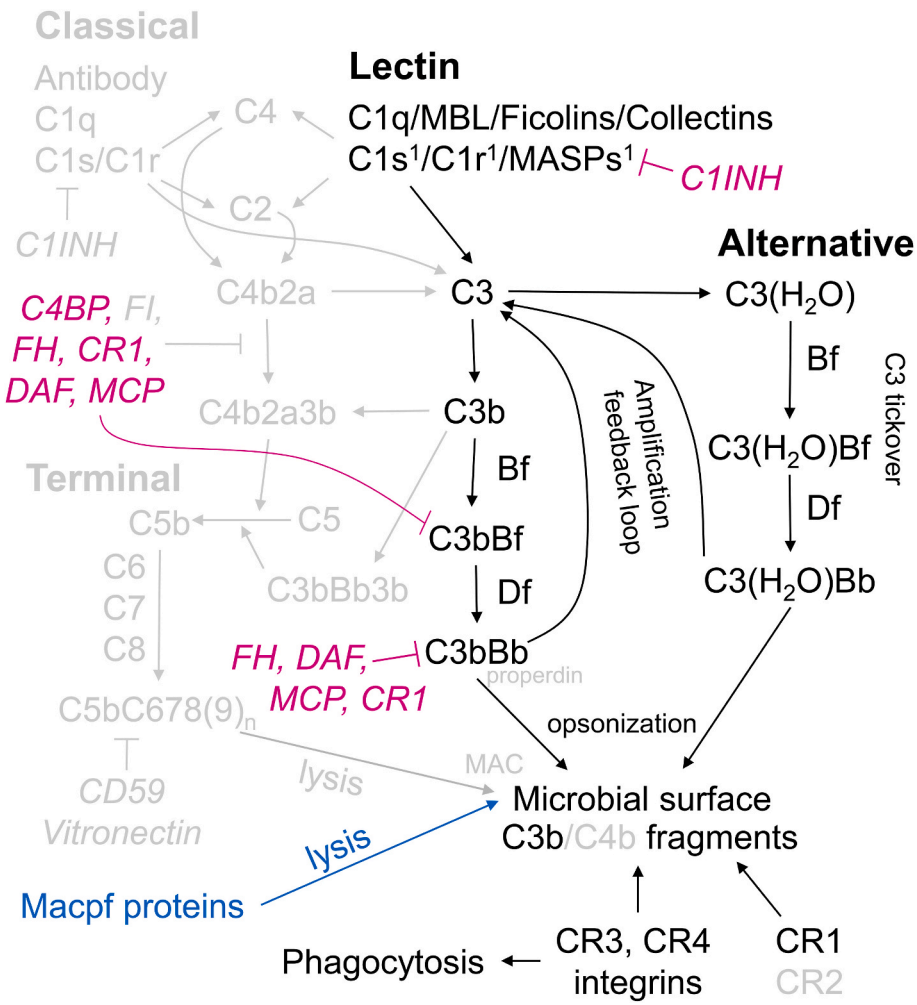



**Fig. 14.** The sea urchin complement system based on searches of deduced proteins encoded in the genome. This diagram of the echinoid complement system is based on the mammalian system (Fig. 1) with unidentified components and activities indicated in gray. The expanded family of SpMacpf proteins (blue) may have anti-pathogen lytic functions that may or may not be associated with the echinoid complement system. See Table 1 for definitions of protein abbreviations. [1]Proteins encoded in asteroid genomes but not in echinoids.

lectin activation pathway. The alternative pathway in *S. purpuratus* may be intact, although verifying the putative function(s) of the SpDf protein(s) will likely require biochemical evaluation for cleaving one or more of the SpBf proteins as has been demonstrated for HsDf (Volanakis and Narayana, 1996). Both the lectin and alternative pathways and the activation of one or both SpC3 proteins suggest that opsonization is an essential and perhaps the only anti-pathogen function of the complement cascade in echinoids. This is consistent with the presence of multiple α and β integrin chains that may be an expansion of CR3-like and CR4-like analogues that function in phagocytosis of opsonized foreign cells and particles by a subset of highly phagocytic coelomocytes. A major difference between the vertebrate and echinoid complement system is the absence of the terminal pathway (Fig. 14). Echinoderm analogues of C5 through C9 are not identified and neither are analogues of the regulatory proteins, CD59 and vitronectin, that, in mammals, regulate the terminal pathway. Although this suggests that the echinoid complement system does not produce MAC for pathogen killing or lysis, the expansion of the SpMacpf-containing proteins in *S. purpuratus* suggests that pathogen killing may employ a different mechanism that may or may not be linked to complement activation.

### *10.1. Potential function of Macpf-containing proteins beyond complement*

The extreme diversification of genes encoding Macpf-containing proteins across the tree of life has led to the hypothesis that ancestral proteins with Macpf domains conferred the capacity to remodel membranes perhaps for fundamental developmental or metabolic functions, and that this remodeling was eventually co-opted for attack and defense in complement and cytotoxic pathways (Anderluh et al., 2014). The proteins in *S. purpuratus* with Macpf domains share a consensus sequence of Y/W-**G**-T-H-F/Y-$X_6$-**G**-G, which is the most conserved motif in the domain and is present across all phyla (Anderluh et al., 2014). The highly conserved glycine residues (in bold) are essential for the functions of proteins with Macpf domains because they enable the conformational changes necessary for insertion into lipid membranes (Gilbert et al., 2013; Hadders et al., 2007).

The SpMacpf proteins may have a variety of immune functions related to membrane remodeling and it is unknown whether they are independent of the complement pathway. Some of the SpMapcf proteins may have functions more analogous to perforin than to the terminal complement components. In mammalian cytotoxic T lymphocytes and NK cells, the perforin-1 protein that is composed of an SP, a Macpf domain, and a unique protein kinase C conserved region 2 (C2) domain (Fig. 7F), is exocytosed from the cytotoxic cell directly onto the target cell to initiate apoptosis (McCormack et al., 2013). In this scenario, extracellular calcium ions catalyze the perforin Macpf domains to form a pore in the target cell membrane, which enables other cytotoxic effectors to be transported into the target cell cytoplasm (Baran et al., 2009; Law et al., 2010; Lopez et al., 2013). Although perforin-1, and by extension, the exocytosis pathway, is exclusive to jawed vertebrates (D'Angelo et al., 2012), other cytotoxic-mediated immunity has been demonstrated in invertebrate chordates that have genes encoding C6-like proteins with Macpf domains (Fig. 7A) (Nonaka, 2014; Suzuki et al., 2002; Wakoh et al., 2004). Interestingly, although there is only a

single perforin gene in mammals, lineage-specific expansions have occurred in frogs and fish (D'Angelo et al., 2012), suggesting that divergent perforin isoforms may indeed exist and function in deuterostome immune systems. The function and mode of action of these vertebrate proteins will require elucidation that may help define the functions of the echinoid analogues with regard to the complement and immune system.

### 10.2. The complement regulatory system in echinoids

The sequences obtained from protein searches and the comparisons among proteins, transcripts, and genes are only the first step in identifying and organizing the multiple sequences in the databases that encode complement regulatory proteins with multiple CCP modules. Currently, the benefits of the multiple transcriptional variants and/or genes encoding CCP-containing regulatory proteins with similar domain structure, but different sequences are not understood. We hypothesize that at least some of these proteins function in regulating the echinoderm complement system. When a host contacts a foreign cell or a foreign tissue and an SpC3 or other SpTecp is activated, controlling opsonization by covalent bond formation from the thioester site that augments phagocytosis (Clow et al., 2004) is required and is likely based on the activity of one or more of the CCP-containing regulatory proteins. The presence of complement regulatory proteins mounted on host cell surfaces is essential for protecting self cells from complement attack, whereas their absence on the surface of foreign cells and particles results in opsonization from thioester bond formation leading to phagocytosis and destruction by phagocytes. Future efforts towards better understanding of the nature of the echinoid regulatory proteins, including when and where they are expressed, and their functions in the complement system, will require biochemical analyses of the regulatory system that protects the host from auto-attack.

### 10.3. Conclusion

For many of the deduced proteins that are reported here, we have only taken the first step in their identification through protein and domain BLAST searches followed by domain analysis of the individual sequences. For a few, more information is available on expression patterns and initial analyses of function. This overview of the genes and encoded proteins in echinoderm genomes, with a focus on the purple sea urchin, has resulted in a hypothetical outline of the structure and putative functions of the echinoderm complement system (Fig. 14), which may be used by others as a basis for addressing the functions of these proteins. In some cases, our results suggest that the echinoderm complement system may vary among the different classes. Commonalities show that C3 homologues are key to the system, which appears to be true for all animal phyla, that the lectin and alternative pathways are essential, that the terminal pathway may only function in vertebrates, and that the echinoderms may employ the expanded Macpf-containing proteins as an alternative approach for pathogen lysis.

## Data availability

Data will be made available on request.

## Acknowledgements

This work was supported by funding from the US National Science Foundation (IOS 1855747) to LCS, an EMBO award (ASTF 353–2016) to NF, and an NIH NRSA award (F32ES029843) to CSS.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dci.2022.104584.

## References

Acton, R.T., Weinheimer, P.F., Hildemann, W.H., Evans, E.E., 1971. Bactericidal antibody response in the Pacific hagfish, *Eptatretus stoutii*. Infect. Immun. 4, 160–166.
Al-Sharif, W.Z., Sunyer, J.O., Lambris, J.D., Smith, L.C., 1998. Sea urchin coelomocytes specifically express a homologue of the complement component C3. J. Immunol. 160, 2983–2997.
Anderluh, G., Kisovec, M., Krasevec, N., Gilbert, J.C., 2014. Distribution of MACPF/CDC proteins. In: Anderluh, G., Kisovec, M., Krasevec, N. (Eds.), MACPF/CDC Proteins - Agents of Defence, Attack and Invasion. Springer Science+Business Media, Dordrecht, pp. 7–30.
Arshinoff, B.I., Cary, G.A., Karimi, K., Foley, S., Agalakov, S., Delgado, F., Lotay, V.S., Ku, C.J., Pells, T.J., Beatman, T.R., Kim, E., Cameron, R.A., Vize, P.D., Telmer, C.A., Croce, J., Ettensohn, C.A., Hinman, V.F., 2022. Echinobase: leveraging an extant model organism database to build a knowledgebase supporting research on the genomics and biology of echinoderms. Nucleic Acids Res. 50, D970–D979.
Baran, K., Dunstone, M., Chia, J., Ciccone, A., Browne, K.A., Clarke, C.J.P., Lukoyanova, N., Saibil, H., Whisstock, J.C., Voskoboinik, I., Trapani, J.A., 2009. The molecular basis for perforin oligomerization and transmembrane pore assembly. Immunity 30, 684–695.
Barratt, J., Weitz, I., 2021. Complement factor D as a strategic target for regulating the alternative complement pathway. Front. Immunol. 12, 712572.
Bertheussen, K., 1981. Endocytosis by echinoid phagocytosis in vitro. I. Recognition of foreign matter. Dev. Comp. Immunol. 5, 241–250.
Bertheussen, K., 1982a. Receptors for complement on echinoid phagocytes. I. The opsonic effect of vertebrate sera on echinoid phagocytosis. Dev. Comp. Immunol. 6, 635–642.
Bertheussen, K., 1982b. Receptors for complement on echinoid phagocytes. II. Purified human complement mediates echinoid phagocytosis. Dev. Comp. Immunol. 6, 635–642.
Bertheussen, K., Seljelid, R., 1982. Receptors for complement on echinoid phagocytes. II. The opsonic effect of vertebrate sera on echinoid phagocytosis. Dev. Comp. Immunol. 6, 423–431.
Bohlson, S.S., Fraser, D.A., Tenner, A.J., 2007. Complement proteins C1q and MBL are pattern recognition molecules that signal immediate and long-term protective immune functions. Mol. Immunol. 44, 33–43.
Bos, I.G.A., Hack, C.E., Abrahams, J.P., 2002. Structural and functional aspects of C1-inhibitor. Immunobiology 205, 518–533.
Buckley, K.M., Ho, E.C.H., Hibino, T., Schrankel, C.S., Shuh, N.W., Wang, G., Rast, J.P., 2017. IL17 factors are early regulators in the gut epithelium during inflammatory response to *Vibrio* in the sea urchin larva. eLife 6, e23481.
Buckley, K.M., Rast, J.P., 2012. Dynamic evolution of Toll-like recetor multigene families in echinoderms. Front. Immunol. 3, 136.
Campbell, R.D., Bentley, D.R., Morley, B.J., 1984. The factor B and C2 genes. Phil. Trans. Roy. Soc. Lond. B 306, 367–378.
Carter, R.H., Spycher, M.O., Ng, Y.C., Hoffman, R., Fearon, D.T., 1988. Synergistic interaction between complement receptor type 2 and membrane IgM on B lymphocytes. J. Immunol. 141, 457–463.
Chou, H.-Y., Lun, C.M., Smith, L.C., 2018. The SpTransformer proteins from the purple sea urchin opsonize bacteria, augment phagocytosis, and retard bacterial growth. PLoS One 13, e0196890.
Clow, L.A., Gross, P.S., Shih, C.-S., Smith, L.C., 2000. Expression of SpC3, the sea urchin complement component, in response to lipopolysaccharide. Immunogenetics 51, 1021–1033.
Clow, L.A., Raftos, D.A., Gross, P.S., Smith, L.C., 2004. The sea urchin complement homologue, SpC3, functions as an opsonin. J. Exp. Biol. 207, 2147–2155.
Coffaro, K.A., 1979. Transplantation immunity in the sea urchin. PhD Dissertation. University of California, Santa Cruz, Santa Cruz, CA.
Coffaro, K.A., 1980. Memory and specificity in the sea urchin Lytechinus pictus. In: Manning, M.J. (Ed.), Phylogeny of Immunological Memory; Developments in Immunology, 10. Elsevier/North Holland Biomedical Press, New York, NY, pp. 77–80.
Coffaro, K.A., Hinegardner, R.T., 1977. Immune response in the sea urchin *Lytechinus pictus*. Science 197, 1389–1390.
Colten, H.R., 1985. Structure organization and expression of the major histocompatibility class III genes. Ann. N. Y. Acad. Sci. 458, 269–276.
D'Angelo, M.E., Dunstone, M.A., Whisstock, J.C., Trapani, J.A., Bird, P.I., 2012. Perforin evolved from a gene duplication of MPEG1, followed by a complex pattern of gene gain and loss within Euteleostomi. BMC Evol. Biol. 12, 59.
Dheilly, N.M., Haynes, P.A., Bove, U., Nair, S.V., Raftos, D.A., 2011. Comparative proteomic analysis of a sea urchin (*Heliocidaris erythrogramma*) antibacteria response revealed in the involvement of apextrin and calreticulin. J. Invertebr. Pathol. 106, 223–229.
Dheilly, N.M., Haynes, P.A., Raftos, D.A., Nair, S.V., 2012. Time course proteomic profiling of cellular responses to immunological challenge in the sea urchin, *Heliocidaris erythrogramma*. Dev. Comp. Immunol. 37, 243–256.
Dishaw, L.J., Smith, S.L., Bigger, C.H., 2005. Characterization of a C3-like cDNA in a coral: phylogenetic implications. Immunogenetics 57, 535–548.
Dupuy, A.G., Caron, E., 2008. Integrin-dependent phagocytosis: spreading from microadhesion to new concepts. J. Cell Sci. 212, 1773–1783.
Ehrlich, P., Morgenroth, J., 1899. Ueber haimolysine. Berliner Klinische Wochenschrift 36, 481–486.
Endo, Y., Matsushita, M., Fujita, T., 2011. The role of ficolins in the lectin pathway of innate immunity. Int. J. Biochem. Cell Biol. 43, 705–712.

Fearon, D.T., 1980. Identification of the membrane glycoprotein that is the C3b receptor of the human erythrocyte, polymorphonuclear leukocyte, B lymphocyte and monocyte. J. Exp. Med. 152, 20–30.
Fujita, T., 2002. Evolution of the lectin-complement pathway and its role in innate immunity. Nat. Rev. Immunol. 2, 346–353.
Fujita, T., Endo, D., Nonaka, M., 2004a. Primitive complement system - recognition and activation. Mol. Immunol. 41, 103–111.
Fujita, T., Matsushita, M., Endo, Y., 2004b. The lectin-complement pathway - its role in innate immunity and evolution. Immunol. Rev. 198, 185–202.
Gaboriaud, C., Frachet, P., Thielens, N.M., Arlaud, G.J., 2012. The human C1q globular domain: structure and recognition of non-immune self ligands. Front. Immunol. 2, 92.
Gilbert, H.E., Asokan, R., Holers, V.M., Perkins, S.J., 2006. The 15 SCR flexible extracellular domains of human complement receptor type 2 can mediate multiple ligand and antigen interactions. J. Mol. Biol. 362, 1132–1147.
Gilbert, R.J.C., Mikelj, M., Dalla Serra, M., Froelich, C.J., Anderluh, G., 2013. Effects of MACPF/CDC proteins on lipid membranes. Cell. Mol. Life Sci. 70, 2083–2098.
Golconda, P., Buckley, K.M., Reynolds, C., Romanello, J., Smith, L.C., 2019. The axial organ and the pharynx are sites of hematopoiesis in the sea urchin. Front. Immunol. 10, 870.
Goldberger, G., Arnaout, M.A., Aden, D., Kay, R., Rits, M., Colten, H.R., 1984. Biosynthesis and postsynthetic processing of human C3b/C4b inactibator (factor I) in three hepatoma cell lines. J. Biol. Chem. 259, 6492–6499.
Gross, P.S., Clow, L.A., Smith, L.C., 2000. SpC3, the complement homologue from the purple sea urchin, *Strongylocentrotus purpuratus*, is expressed in two subpopulations of the phagocytic coelomocytes. Immunogenetics 51, 1034–1044.
Groves, E., Dart, A., Covarelli, V., Caron, E., 2008. Molecular mechanisms of phagocytic uptake by mammalian cells. Cell. Mol. Life Sci. 65, 1957–1976.
Haag, E.S., Sly, B.J., Andrews, M.E., Raff, R.A., 1999. Apextrin, a novel extracellular protein associated with larval ectoderm evolution in *Heliocidaris erythrogramma*. Dev. Biol. 211, 77–87.
Hadders, M.A., Beringer, D.X., Gros, P., 2007. Structure of C8alpha-MACPF reveals mechanism of membrane attack in complement immune defense. Science 317, 1552–1554.
Hakansson, K., Reid, K.B.M., 2000. Collectin structure: a review. Protein Sci. 9, 1607–1617.
Hanley, P.J., Hook, J.W., Raftos, D.A., Gooley, A.A., Trent, R., Raison, R.L., 1992. Hagfish humoral defense protein exhibits structural and functional homology with mammalian complement components. Proc. Natl. Acad. Sci. USA 89, 7910–7914.
Hansen, S.W., Ohtani, K., Suzuki, Y., Wakamiya, N., 2016. The collectins CL-L1, Cl-K1 and ClL-P1 and their roles in complement and innate immunity. Immunobiology 221, 1058–1067.
Hellwage, J., Eberle, F., Babuke, T., Seeberger, H., Richter, H., Kunert, A., Härtl, A., Zipfel, P.F., Jokiranta, T.S., Jozsi, M., 2006. Two factor H-related proteins from the mouse: expression analysis and functional characterization. Immunogenetics 58, 883–893.
Hibino, T., Loza-Coll, M., Messier, C., Majeske, A.J., Cohen, A., Terwilliger, D.P., Buckley, K.M., Brockton, V., Nair, S., Berney, K., Fugmann, S.D., Anderson, M.K., Pancer, Z., Cameron, R.A., Smith, L.C., Rast, J.P., 2006. The immune gene repertoire encoded in the purple sea urchin genome. Dev. Biol. 300, 349–365.
Hildemann, W.H., Dix, T.G., 1972. Transplantation reactions of tropical Australian echinoderms. Transplantation 14, 624–633.
Ho, E.C.H., Buckley, K.M., Schrankel, C.S., Schuh, N.W., Hibino, T., Solek, C.M., Bae, K., Wang, G., Rast, J.P., 2016. Perturbation of gut bacteria induces a coordinated cellular immune response in the purple sea urchin larva. Immunol. Cell Biol. 94, 861–874.
Hourcade, D.E., Mitchell, L.M., 2011. Access to the complement factor B scissile bond is facilitated by association of factor B with C3b protein. J. Biol. Chem. 286, 35725–35732.
Huang, G., Huang, S., Yan, X., Yang, P., Li, J., Xu, W., Zhang, L., Wang, R., Yu, Y., Yuan, S., Chen, S., Luo, G., Xu, A., 2014. Two apextrin-like proteins mediate extracellular and intracellular bacterial recognition in amphioxus. Proc. Natl. Acad. Sci. USA 111, 13469–13474.
Huang, S., Yuan, S., Guo, L., Yu, Y., Li, J., Wu, T., Liu, T., Yang, M., Wu, K., Liu, H., Ge, J., Yu, Y., Huang, H., Dong, M., Yu, C., Chen, S., Xu, A., 2008. Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. Genome Res. 18, 1112–1126.
Ishiguro, H., Kobayashi, K., Suzuki, M., Titani, K., Tomonaga, S., Kurosawa, Y., 1992. Isolation of a hagfish gene that encodes a complement component. EMBO J. 11, 829–837.
Józsi, M., Richter, H., Löschmann, I., Skerka, C., Buck, F., Beisiegel, U., Erdei, A., Zipfel, P.F., 2005. FHR-4A: a new factor H-related protein is encoded by the human *FHR-4* gene. Eur. J. Hum. Genet. 13, 321–329.
Kaplan, G., Bertheussen, K., 1977. The morphology of echinoid phagocytes and mouse peritoneal macrophages during phagocytosis in vitro. Scand. J. Immunol. 6, 1289–1296.
Karp, R.D., Hildemann, W.H., 1976. Specific allograft reactivity in the sea star *Dermasterias imbricata*. Transplantation 22, 434–439.
Keshi, H., Sakamoto, T., Kawai, T., Ohtani, K., Katoh, T., Jang, S.-J., Motomura, W., Yoshizaki, T., Fukuda, M., Koyama, S., Fukuzawa, J., Fukuoh, A., Yoshida, I., Suzuki, Y., Wakamiya, N., 2006. Identification and characterization of a novel human collectin CL-K1. Microbiol. Immunol. 50, 1001–1013.
Kim, D.D., Song, W.-C., 2006. Membrane complement regulatory proteins. Clin. Immunol. 118, 127–136.
Kimura, A., Sakaguchi, E., Nonaka, M., 2009. Multi-component complement system of Cnidaria: C3, Bf, and MASP genes expressed in the endodermal tissues of a sea anemone, *Nematostella vectensis*. Immunobiology 214, 165–178.
Kolb, W.P., Muller-Eberhard, H.J., 1975. The membrane attack mechanism of complement. J. Exp. Med. 141, 724–735.
Krych-Goldberg, M., Atkinson, J., 2001. Structure-function relationships of complement receptor type 1. Immunol. Rev. 180, 112–122.
Lachmann, P.J., 1979. An evolutionary view of the complement system. Behring Institute Mitteilungen 63, 25–37.
Lachmann, P.J., 2009. The amplification loop of the complement pathways. Adv. Immunol. 104, 115–149.
Law, R.H.P., Lukoyanova, N., Voskoboinik, I., Caradoc-Davies, T.T., Baran, K., Dunstone, M.A., D'Angelo, M.E., Orlova, E.V., Coulibaly, F., Verschoor, S., Browne, K.A., Ciccone, A., Kuiper, M.J., Bird, P.I., Trapani, J.A., Saibil, H.R., Whisstock, J.C., 2010. The structural basis for membrane binding and pore formation by lymphocyte perforin. Nature 468, 447–451.
Leonard, L.A., Strandberg, J.D., Winkelstein, J.A., 1990. Complement-like activity in the sea star, *Asterias forbesi*. Dev. Comp. Immunol. 14, 19–30.
Li, J., Li, Y., Fan, Z., Chen, S., Yan, X., Yue, Z., Huang, G., Liu, S., Zhang, H., Chen, S., Dong, M., Xu, A., Huang, S., 2021a. Two amphioxus ApeC-containing proteins bind to microbes and inhibit the TRAF6 pathway. Front. Immunol. 12, 715245.
Li, Y., Li, J., Yan, X., Chen, S., Wu, C., Huang, H., Shi, Y., Huang, G., Dong, M., Xu, A., Huang, S., 2021b. Broad distribution, high diversity and ancient origin of the ApeC-containing proteins. Mol. Phylogenet. Evol. 155, 107009.
Liu, Y., Endo, D., Iwake, M., Nakata, M., Matsushita, M., Wada, I., Inoue, K., Munakata, M., Fujita, T., 2005. Human M-ficolin is a secretory protein that activates the lectin complement pathway. J. Immunol. 175, 3150–3156.
Lopez, J.A., Jenkins, M.R., Rudd-Schmidt, J.A., Brennan, A.J., Danne, J.C., Mannering, S. I., Trapani, J.A., Voskoboinik, I., 2013. Rapid and unidirectional perforin pore delivery at the cytotoxic immune synapse. J. Immunol. 191, 2328–2334.
Lowrey, D.M., Aebischer, T., Olsen, K., Lichtenheld, M., Rupp, F., Hengartner, H., Podack, E.R., 1989. Cloning, analysis, and expression of murine perforin 1 cDNA, a component of cytolytic T-cell granules with homology to complement component C9. Proc. Natl. Acad. Sci. USA 86, 247–251.
Lu, J., Thiel, S., Wiedemann, H., Timpl, R., Reid, K.B.M., 1990. Binding of the pentamer/hexamer forms of mannan-binding protein to zymosan activates the proenzyme $C1r_2C1s_2$ complex, of the classical pathway of complement, without involvement of C1q. J. Immunol. 144, 2287–2294.
Majeske, A.J., Oleksyk, T.K., Smith, L.C., 2013. The *Sp185/333* immune response genes and proteins are expressed in cells dispersed within all major organs of the adult purple sea urchin. Innate Immun. 19, 569–587.
Marchalonis, J.J., 1977. Immunity in Evolution. Harvard University Press, Cambridge, MA.
Marchalonis, J.J., Edelman, G.M., 1968. Phylogenetic origins of antibody structure: III. Antibodies in the primary immune response of the sea lamprey, *Petromyzon marinus*. J. Exp. Med. 127, 891–914.
Marsden, M., Burke, R.D., 1997. Cloning and characterization of novel beta integrin subunits from a sea urchin. Dev. Biol. 181, 234–245.
Matsushita, M., Fujii, T., 1992. Activation of the classical complement pathway by mannose-binding protein in association with a novel C1s-like serine protease. J. Exp. Med. 176, 1497–1502.
Matsushita, M., Fujita, T., 2001. Ficolins and the lectin complement pathway. Immunol. Rev. 180, 78–85.
Matsushita, M., Fujita, T., 2002. The role of ficolins in innate immunity. Immunobiology 205, 490–497.
Matz, H., Munir, D., Logue, J., Dooley, H., 2021. The immunoglobulins of cartilaginous fishes. Dev. Comp. Immunol. 115, 103873.
Mayer, M.M., 1973. The complement system. Sci. Am. 229, 54–69.
McCormack, R., De Armas, L., Shiratsuchi, M., Podack, E.R., 2013. Killing machines: three pore-forming proteins of the immune system. Immunol. Res. 57, 268–278.
Meri, S., Morgan, B.P., Davies, A., Daniels, R.H., Olavesen, M.G., Waldmann, H., Lachmann, P.J., 1990. Human protectin (CD59), an 18,000-20,000 MW complement lysis restricting factor, inhibits C5b-8 catalysed insertion of C9 into lipid bilayers. Immunology 71, 1–9.
Merle, N.S., Church, S.E., Fremeaux-Bacchi, V., Roumenina, L.T., 2015a. Complement system part I - molecular mechanisms of activation and regulation. Front. Immunol. 6, 262.
Merle, N.S., Noe, R., Halbwachs-Mecarelli, L., Fremeaux-Bacchi, V., Roumenina, L.T., 2015b. Complement system part II: role in immunity. Front. Immunol. 6, 257.
Merle, N.S., Singh, P., Rahman, J., Kemper, C., 2021. Integrins meet complement: the evolutionary tip of an iceberg orchestrating metabolism and immunity. Br. J. Pharmacol. 178, 2754–2770.
Miyashita, T., Coates, M.I., Farrar, R., Larson, P., Manning, P.L., Wogelius, R.A., Edwards, N.P., Anne, J., Bermann, U., R, P.A., Currie, P.J., 2019. Hagfish from the Cretaceous Tethys Sea and a reconciliation of the morphological-molecular conflict in early vertebrate phylogeny. Proc. Natl. Acad. Sci. USA 116, 2146–2151.
Murray, G., Reed, C., Marsden, M., Rise, M., Wang, D., Burke, R.D., 2000. The alphaB betaC integrin is expressed on the surface of the sea urchin egg and removed at fertilization. Dev. Biol. 227, 633–647.
Nakao, M., Matsumoto, M., Nakazawa, M., Fujiki, K., Yano, T., 2002. Diversity of complement factor B/C2 in the common carp (*Cyprinus carpio*): three isotypes of B/C2-A expressed in different tissues. Dev. Comp. Immunol. 26, 533–541.
Nonaka, M., 2014. Evolution of the complement system. Subcell. Biochem. 80, 31–43.
Nonaka, M., Azumi, K., 1999. Opsonic complement system of the solitary ascidian, *Halocynthia roretzi*. Dev. Comp. Immunolol. 23, 421–427.

Nonaka, M., Fujii, T., Kaidoh, T., Natsuume-Sakai, S., Nonaka, M., Yamaguchi, N., Takahashi, M., 1984. Purification of a lamprey complement protein homologous to the third component of the mammalian complement system. J. Immunol. 133, 3242–3249.
Nonaka, M., Yoshizaki, F., 2004. Primitive complement system of invertebrates. Immunol. Rev. 198, 203–215.
Ohta, M., Ito, H., Masuda, K., Tanaka, S., Arakawa, Y., Wacharotayankun, R., Kato, N., 1992. Mechanisms of antibacterial action of tachyplesins and polyphemusins, a group of antimicrobial peptides isolated from horseshoe crab hemocytes. Antimicrob. Agents Chemother. 36, 1460–1465.
Ohtani, K., Suzuki, Y., Wakamiya, N., 2012. Biological functions of the novel collectins CL-L1, Cl-K1, Cl-P1. J. Biomed. Biotech., 493945
Pettinello, R., Dooley, H., 2014. The immunoglobulins of cold-blooded vertebrates. Biomolecules 4, 1045–1069.
Pillemer, L., Blum, L., Lepow, I.H., Ross, O.A., Todd, E.W., Wardlaw, A.C., 1954. The properdin system and immunity. I. Demonstration and isolation of a new serum protein, peroperdin, and its role in immune phenomena. Science 120, 279–285.
Polgar, L., 2005. The catalytic triad of serine peptidases. Cell. Mol. Life Sci. 62, 2161–2172.
Pouw, R.B., Vredevoogd, D.W., Kuijpers, T.W., Wouters, D., 2015. Of mice and men: the factor H protein family and complement regulation. Mol. Immunol. 67, 12–20.
Preissner, K.T., 1991. Structure and biological role of vitronectin. Annu. Rev. Cell Biol. 7, 275–310.
Raison, R.L., Coverley, J., Hook, J.W., Towns, P., Weston, K.M., Raftos, D.A., 1994. A cell-surface opsonic receptor on leucocytes from the phylogenetically primitive vertebrate, *Eptatretus stouti*. Immunol. Cell Biol. 72, 326–332.
Raison, R.L., Hull, C.J., Hildemann, W.H., 1978a. Characterization of immunoglobulin from the Pacific hagfish, a primitive vertebrate. Proc. Natl. Acad. Sci. USA 75, 5679–5682.
Raison, R.L., Hull, C.J., Hildemann, W.H., 1978b. Production and specificity of antibodies to streptococci in the Pacific hafgish, *Eptatretus stoutii*. Dev. Comp. Immunol. 2, 253–261.
Rast, J.P., Smith, L.C., Loza-Coll, M., Hibino, T., Litman, G.W., 2006. Genomic insights into the immune system of the sea urchin. Science 314, 952–956.
Reid, K.B.M., 2018. Complement component C1q: historical perspective of a functionally versatile, and structurally unusual, serum protein. Front. Immunol. 9, 764.
Reid, K.B.M., Porter, R.R., 1981. The proteolytic activtion systems of complement. Annu. Rev. Biochem. 50, 433–464.
Romer, A.S., 1967. Major steps in vertebrate evolution. Science 158, 1629–1637.
Ross, G.D., Lambris, J.D., Cain, J.A., Newman, S.L., 1982. Generation of three different fragments of bound C3 with purified factor I or serum. I. Requirements for factor H vs. CR1 cofactor activity. J. Immunol. 129, 2051–2060.
Sackstein, R., Colten, H.R., 1984. Molecular regulation of MHC class III (C4 and fctor B) gene expression in mouse peritoneal macrophages. J. Immunol. 133, 1618–1626.
Sarma, J.V., Ward, P.A., 2011. The complement system. Cell Tissue Res. 343, 227–235.
Schrankel, C.S., 2017. Gene regulatory control of immune cell specification and differentiation in the sea urchin embryo and larva. PhD Dissertation. University of Toronto, Toronto Canada.
Schwaeble, W., Dahl, M.R., Thiel, S., Stover, C., Jensenius, J.C., 2002. The mannan-binding lectin-associated serine proteaes (MASPs) and Map19: four components of the lectin activtion complex encoded by two genes. Immunobiology 205, 455–466.
Shah, M., Brown, K.M., Smith, L.C., 2003. The gene encoding the sea urchin complement protein, SpC3, is expressed in embryos and can be upregulated by bacteria. Dev. Comp. Immunol. 27, 529–538.
Shinkai, Y., Takio, K., Okumura, K., 1988. Homology of perforin to the ninth component of complement (C9). Nature 344, 525–527.
Skerka, C., Chen, Q., Fremeaux-Bacchi, V., T, R.L., 2013. Complement factor H related proteins (CFHRs). Mol. Immunol. 56, 170–180.
Smith, L.C., 2002. Thioester function is conserved in SpC3, the sea urchin homologue of the complement component C3. Dev. Comp. Immunol. 26, 603–614.
Smith, L.C., Azumi, K., Nonaka, M., 1999. Complement systems in invertebrates. The ancient alternative and lectin pathways. Immunopharmacology 42, 107–120.
Smith, L.C., Chang, L., Britten, R.J., Davidson, E.H., 1996. 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. J. Immunol. 156, 593–602.
Smith, L.C., Clow, L.A., Terwilliger, D.P., 2001. The ancestral complement system in sea urchins. Immunol. Rev. 180, 16–34.
Smith, L.C., Davidson, E.H., 1992. The echinoid immune system and the phylogenetic occurrence of immune mechanisms in deuterostomes. Immunol. Today 13, 356–362.
Smith, L.C., Rast, J.P., Brockton, V., Terwilliger, D.P., Nair, S.V., Buckley, K.M., Majeske, A.J., 2006. The sea urchin immune system. Invertebr. Surviv. J. 3, 25–39.
Smith, L.C., Shih, C.S., Dachenhausen, S.G., 1998. Coelomocytes express SpBf, a homologue of factor B, the second component in the sea urchin complement system. J. Immunol. 161, 6784–6793.
Sodergren, E., Weinstock, G.M., Davidson, E.H., Cameron, R.A., Gibbs, R.A., Angerer, R. C., Angerer, L.M., Arnone, M.I., Burgess, D.R., Burke, R.D., Coffman, J.A., Dean, M., Elphick, M.R., Ettensohn, C.A., Foltz, K.R., Hamdoun, A., Hynes, R.O., Klein, W.H., Marzluff, W., McClay, D.R., Morris, R.L., Mushegian, A., Rast, J.P., Smith, L.C., Thorndyke, M.C., Vacquier, V.D., Wessel, G.M., Wray, G., Zhang, L., Elsik, C.G., Ermolaeva, O., Hlavina, W., Hofmann, G., Kitts, P., Landrum, M.J., Mackey, A.J., Maglott, D., Panopoulou, G., Poustka, A.J., Pruitt, K., Sapojnikov, V., Song, X., Souvorov, A., Solovyev, V., Wei, Z., Whittaker, C.A., Worley, K., Durbin, K.J., Shen, Y., Fedrigo, O., Garfield, D., Haygood, R., Primus, A., Satija, R., Severson, T., Gonzalez-Garay, M.L., Jackson, A.R., Milosavljevic, A., Tong, M., Killian, C.E., Livingston, B.T., Wilt, F.H., Adams, N., Belle, R., Carbonneau, S., Cheung, R., Cormier, P., Cosson, B., Croce, J., Fernandez-Guerra, A., Geneviere, A.M., Goel, M., Kelkar, H., Morales, J., Mulner-Lorillon, O., Robertson, A.J., Goldstone, J.V., Cole, B., Epel, D., Gold, B., Hahn, M.E., Howard-Ashby, M., Scally, M., Stegeman, J. J., Allgood, E.L., Cool, J., Judkins, K.M., McCafferty, S.S., Musante, A.M., Obar, R.A., Rawson, A.P., Rossetti, B.J., Gibbons, I.R., Hoffman, M.P., Leone, A., Istrail, S., Materna, S.C., Samanta, M.P., Stolc, V., Tongprasit, W., Tu, Q., Bergeron, K.F., Brandhorst, B.P., Whittle, J., Berney, K., Bottjer, D.J., Calestani, C., Peterson, K., Chow, E., Yuan, Q.A., Elhaik, E., Graur, D., Reese, J.T., Bosdet, I., Heesun, S., Marra, M.A., Schein, J., Anderson, M.K., Brockton, V., Buckley, K.M., Cohen, A.H., Fugmann, S.D., Hibino, T., Loza-Coll, M., Majeske, A.J., Messier, C., Nair, S.V., Pancer, Z., Terwilliger, D.P., Agca, C., Arboleda, E., Chen, N., Churcher, A.M., Hallbook, F., Humphrey, G.W., Idris, M.M., Kiyama, T., Liang, S., Mellott, D., Mu, X., Murray, G., Olinski, R.P., Raible, F., Rowe, M., Taylor, J.S., Tessmar-Raible, K., Wang, D., Wilson, K.H., Yaguchi, S., Gaasterland, T., Galindo, B.E., Gunaratne, H.J., Juliano, C., Kinukawa, M., Moy, G.W., Neill, A.T., Nomura, M., Raisch, M., Reade, A., Roux, M.M., Song, J.L., Su, Y.H., Townley, I.K., Voronina, E., Wong, J.L., Amore, G., Branno, M., Brown, E.R., Cavalieri, V., Duboc, V., Duloquin, L., Flytzanis, C., Gache, C., Lapraz, F., Lepage, T., Locascio, A., Martinez, P., Matassi, G., Matranga, V., Range, R., Rizzo, F., Rottinger, E., Beane, W., Bradham, C., Byrum, C., Glenn, T., Hussain, S., Manning, G., Miranda, E., Thomason, R., Walton, K., Wikramanayke, A., Wu, S.Y., Xu, R., Brown, C.T., Chen, L., Gray, R.F., Lee, P.Y., Nam, J., Oliveri, P., Smith, J., Muzny, D., Bell, S., Chacko, J., Cree, A., Curry, S., Davis, C., Dinh, H., Dugan-Rocha, S., Fowler, J., Gill, R., Hamilton, C., Hernandez, J., Hines, S., Hume, J., Jackson, L., Jolivet, A., Kovar, C., Lee, S., Lewis, L., Miner, G., Morgan, M., Nazareth, L.V., Okwuonu, G., Parker, D., Pu, L.L., Thorn, R., Wright, R., 2006. The genome of the sea urchin, *Strongylocentrotus purpuratus*. Science 314, 941–952.
Sommer, F., Awazu, S., Anton-Erxleben, F., Jiang, D., Klimovich, A.V., Klimovich, B.V., Samoilovich, M.P., Satou, Y., Krüss, M., Gelhaus, C., Kürn, U., Bosch, T.C.G., Khalturin, K., 2012. Blood system formation in the urochordate *Ciona intestinalis* requires the variable receptor vCRL1. Mol. Biol. Evol. 29, 3081–3093.
Sonnen, A.F.-P., Henneke, P., 2014. Structural biology of the membrane attack complex. Subcell. Biochem. 80, 83–116.
Suzuki, M.M., Satoh, N., Nonaka, M., 2002. C6-like and C3-like molecules from the cephalochordate, amphioxus, suggest a cytolytic complement system in invertebrates. J. Mol. Evol. 54, 671–679.
Takahashi, K., Ip, E.W.K., Michelow, I.D., Ezekowitz, R.A.B., 2006. The mannose-binding lectin: a prototypic pattern recognition molecule. Curr. Opin. Immunol. 18, 16–23.
Terwilliger, D.P., Clow, L.A., Gross, P.S., Smith, L.C., 2004. Constitutive expression and alternative splicing of the exons encoding SCRs in Sp152, the sea urchin homologue of complement factor B. Implications on the evolution of the Bf/C2 gene family. Immunogenetics 56, 531–543.
Thielens, N.M., Tedesco, F., Bohlson, S.S., Gaboriaud, C., Tenner, A.J., 2017. C1q: a fresh look upon an old molecule. Mol. Immunol. 89, 73–83.
Tschopp, J., Masson, D., Schafer, S., Peitsch, M.C., Preissner, K.T., 1988. The heparin binding domains of S-protein/vitronectin binds to complement components C7, C8, C9 and perforin from cytolytic T-cells and inhibits their lytic activities. Biochemistry 27, 4103-4009.
van den Berg, C.W., Cinek, T., Hallett, M.B., Horejsi, V., Morgan, B.P., 1995. Exogenous glycosyl phosphatidylinositol-anchored CD59 associates with kinases in membrane clusters on U947 cells and becomes Ca(2+)-signaling competent. J. Cell Biol. 131, 669–677.
Volanakis, J.E., Narayana, S.V.L., 1996. Complement factor D, a novel serine protease. Protein Sci. 5, 553–564.
Vorup-Jensen, T., Jensen, K., 2018. Structural immunology of complement receptors 3 and 4. Front. Immunol. 9, 2716.
Wakoh, T., Ikeda, M., Uchino, R., Azumi, K., Nonaka, M., Kohara, Y., Metoki, H., Satou, Y., Satoh, N., Satake, M., 2004. Identification of transcripts expressed preferentially in hemocytes of *Ciona intestinalis* that can be used as molecular markers. DNA Res. 11, 345–352.
Wallis, R., 2007. Interactions between mannose-binding lectin and MASPs during complement activtion by the lectin pathway. Immunobiology 212, 289–299.
Weis, J.J., Toothaker, L.E., Smith, J.A., Weis, J.H., Fearon, D.T., 1988. Structure of the human B lymphocyte receptor for C3d and the Epstein-Barr virus and relatedness to other members of C3/C4 proteins. J. Exp. Med. 167, 1047–1066.
Whittaker, C.A., Bergeron, K.F., Whittle, J., Brandhorst, B.P., Burke, R.D., Hynes, R.O., 2006. The echinoderm adhesome. Dev. Biol. 300, 252–266.
Zhong, L., Zhang, F., Chang, Y., 2012. Gene cloning and function analysis of complement B factor-2 of *Apostichopus japonicus*. Fish Shellfish Immunol. 33, 504–513.