Ancient Origin of the Complement System: Emerging Invertebrate Models

Maria Rosaria Pinto[1], Daniela Melillo[1], Stefano Giacomelli[1], Georgia Sfyroera[2] and John D. Lambris[2]

[1] Stazione Zoologica "Anton Dohrn", Laboratory of Cell Biology, pinto@szn.it
[2] University of Pennsylvania, Department of Pathology and Laboratory Medicine, lambris@mail.med.upenn.edu

1 Introduction

All metazoans are endowed with defense mechanisms against invading pathogens. The defense machinery has increased in complexity during evolution, with the addition of new and diverse components that have acquired the ability to cooperate with each other to provide an efficient and prompt immune response. The highest level of functionality has been reached in higher vertebrates, in the form of integrated action by the innate and adaptive immune systems.

In this context, the complement system, the major effector arm of vertebrate innate immunity, represents a link between innate and adaptive immunity (Song et al. 2000). In higher vertebrates, it involves more than 30 humoral and cell membrane proteins that are organized into different activation and effector pathways. The effector function of the complement system can be activated through three different activation pathways: the alternative pathway, triggered by the binding of a complement component to the pathogen surface; the mannose-binding lectin (MBL) pathway, initiated by the binding of serum lectins to a pathogen; and the classical pathway, triggered by the binding of an antibody to antigen. All these pathways lead to the activation of C3, the acknowledged molecular pillar of the complement system. Its proteolytic cleavage by C3 convertase triggers the effector function of the complement, leading to the recruitment of inflammatory cells and opsonization of the pathogen, or to its lysis through the formation of the membrane attack complex (MAC) (Lambris 1990).

In the past decade, in the context of the renewed interest in innate immunity, the complement system has been investigated in increasing depth. One successful approach to analyzing the complement system has involved the study of its evolutionary origin. The search for complement components has been carried out in
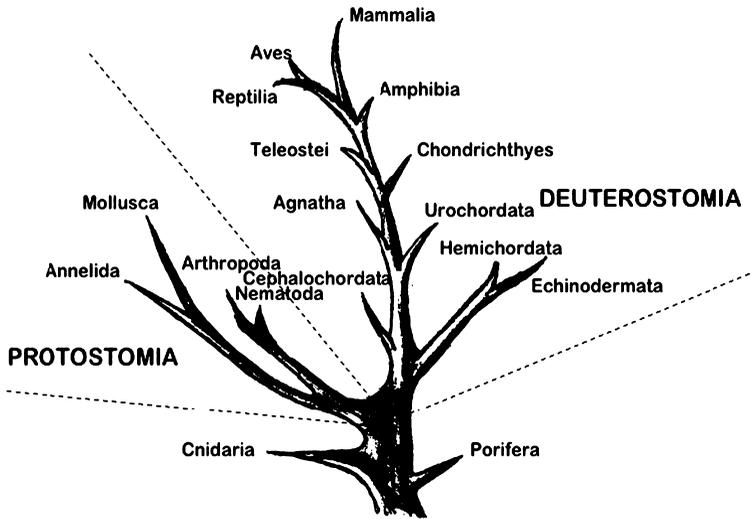
Fig. 1. Phylogenetic tree.

very divergent species, aided by the powerful tools provided by computational biology and the genome projects that are ongoing in many invertebrate species. A surprising result of these endeavors has been the identification of complement genes in very ancient species belonging to the phylum Cnidaria, which have been in existence since about 1,300 million years ago (mya) (Dishaw et al. 2005; Nonaka and Kimura 2006). Several authors have previously reviewed the results of these investigations, focusing their attention mainly on the molecular evolution of the invertebrate complement components. Indeed, data on the biological role of these components are extremely limited because of a number of factors, including the slow development, until recently, of interest in the invertebrate complement system, the paucity of the available biological material, and the relative newness of the available functional assays.

In this review we summarize the published data on invertebrate complement components, giving priority to the functional aspects of these components when known, and analyzing the molecular signatures and domain structures that suggest particular complement-associated activities. The phyla are discussed in the order of their phylogenetic position (Fig. 1); the accession numbers of the cited complement molecules are given in Table I.

Table I. Accession numbers of the amino acid sequences cited in the text

Swiftia exserta C3	AAN86548
Carcinoscorpius rotundicauda C3	AF517564
Strongylocentrotus purpuratus C3	AAC14396
Branchiostoma belcheri C3	BAB47146
Ciona intestinalis C3-1	CAC85959
Ciona intestinalis C3-2	CAC85958
Halocynthia roretzi C3	BAA75069
Carcinoscorpius rotundicauda Bf	AAV65032
Strongylocentrotus purpuratus Bf	AAC79682
Ciona intestinalis Bf-1	BAD89299
Ciona intestinalis Bf-2	BAD89300
Ciona intestinalis Bf-3	BAD89301
Halocynthia roretzi MBL	BAB69891
Halocynthia roretzi ficolin1	BAB60704
Halocynthia roretzi ficolin2	BAB60705
Halocynthia roretzi ficolin3	BAB60706
Halocynthia roretzi ficolin4	BAB60707
Branchiostoma belcheri MASP1	BAC75886
Branchiostoma belcheri MASP3	BAC75887
Halocynthia roretzi MASPa	BAA19762
Halocynthia roretzi MASPb	BAA19763
Branchiostoma belcheri C6	BAB47147
Ciona intestinalis C3aR	CAI84650
Halocynthia roretzi Integrin α Hr1	BAB21479
Halocynthia roretzi Integrin α Hr2	BAB21480
Halocynthia roretzi Integrin β Hr1	BAD15077
Halocynthia roretzi Integrin β Hr2	BAD15078
Strongylocentrotus purpuratus Sp5	AAR87482
Strongylocentrotus purpuratus Sp5013	AAR87483

2 Cnidaria

A search for the presence of complement genes has been carried out in the genome of the sea anemone *Nematostella vectensis*. This organism belongs to the phylum Cnidaria, which diverged from the Bilateria about 1,300 mya, before the protostome/deuterostome divergence, which occurred approximately 1,000 mya. This analysis was based on a search for predicted domain structures, taking into account the fact that at least five complement gene families exhibit a unique domain combination found only among human complement genes. This search produced evidence for the presence of two complement components, C3 and factor B (Bf) (Nonaka and Kimura 2006).

This finding is in agreement with the report of Dishaw and collaborators (2005), who cloned from the gorgonian coral *Swiftia exserta* a full-length cDNA for a C3-like gene, *SeC3*. The analysis of the deduced amino acid sequence of this cDNA revealed a 24 and 45% identity and similarity, respectively, with the human C3 molecule. Many molecular features of mammalian C3, such as the canonical thioester site, the associated catalytic histidine, the β-α cleavage site, and the C3a anaphylatoxin domain, were present in SeC3. The existence of a putative α-γ cleavage site suggested that SeC3, unlike mammalian C3, is a three-chain molecule similar to mammalian C4 (Karp et al. 1981), lamprey C3 (Nonaka 1994), horseshoe crab C3 (Zhu et al. 2005), and cobra venom factor (Vogel et al. 1996). All the canonical cysteine residues of mammalian C3 were conserved in SeC3, with the exception of those bridging the α and β chains. The authors suggested that the β-chain could be either released by the functional protein or associated with the α-chain through different interactions. They also suggested that the absence of the disulfide bond leaves the anaphylatoxin region more exposed to the convertase, facilitating its enzymatic cleavage. This site in SeC3 shows a context for the R-S target bond (RTR-S) that is different from that specific to mammalian C3 (LAR-S). In a phylogenetic analysis, carried out using the minimum evolution distance method on 52 thiol ester proteins (TEPs) from different vertebrate and invertebrate species, SeC3 clustered with the deuterostome invertebrate C3-like proteins, the sister group of the C3, C4 and C5 vertebrate proteins.

3 Arthropoda

The horseshoe crab *Carcinoscorpius rotundicauda* (Arthropoda) can be considered a "living fossil," as it first appeared about 550 mya. Among the protostome species analyzed thus far, *C. rotundicauda* is the only one that possesses complement system components (Zhu et al. 2005). In fact, a search for complement system genes in the available genomes of the protostomes *Anopheles gambiae*, *Drosophila melanogaster*, and *Caenorhabditis elegans* did not produce any results, and this absence has been interpreted as a secondary loss that has occurred many times at various stages during protostome evolution (Nonaka and Kimura 2006).

Using different bacterial matrices, Zhu et al. (2005) have identified several proteolytic fragments of a C3 homolog, CrC3, in the plasma of *C. rotundicauda*. A cDNA encoding the corresponding molecule has been cloned and sequenced. Analysis of the deduced amino acid sequence of CrC3 revealed the presence of all the canonical features of C3, including a domain showing high sequence similarity to that of vertebrate C3/C4/C5 anaphylatoxins. However, the vertebrate C3 convertase cleavage site LXR/S was replaced in CrC3 by EGR/F, which is more similar to the motif QGR/S of *Ciona intestinalis* C3-1 (Marino et al. 2002). A three-chain structure of the mature CrC3 was predicted on the basis of the presence of the C3/C4/C5 β-α cleavage site RKKR and two α-γ processing motifs. This structure, which is characteristic of higher vertebrate C4 molecule, is common to SeC3 and lamprey C3. C3-mediated opsonization was also observed in form of the attachment of CrC3 to the

surfaces of different groups of microbes and by its subsequent proteolytic fragmentation.

Two other major proteins, homologs of the horseshoe crab tachylectins and vertebrate ficolins (Kairies et al. 2001), have been recovered from bacterial surfaces. These two proteins, the carcinolectins CL5a and CL5b, act synergistically in recognizing various pathogens and initiating the activation of the complement system in the lectin pathway (Matsushita and Fujita 2002).

A homolog of vertebrate C2 and Bf (CrC2/Bf) has also been identified in *C. rotundicauda* by subtractive hybridization of hepatopancreas cDNAs from naïve and bacteria-challenged horseshoe crabs. CrC2/Bf showed the highest identity to sea urchin C2/Bf and exhibited the main canonical sequence motifs of all C2/Bf molecules: the factor D cleavage site, Mg^{2+}-binding site, and serine protease catalytic triad (H, D, S), together with the expected overall domain architecture, consisting of multiple complement control protein (CCP) modules, a von Willebrand factor (VWF) domain and a trypsin serine protease domain. In contrast to vertebrate C2/Bf, which contains three CCP modules, CrC2/Bf, like lower deuterostome C2/Bf, contains five CCP modules. It has also been suggested that the Mg^{2+}-dependent serine protease activity found in LPS-treated horseshoe crab plasma can participate in the CrC3 activation that is mediated by CrC2/Bf. This activity is also correlated with the production of a peptide homologous to the vertebrate C3a fragment, whose presence in horseshoe crab plasma was inferred from the comparison of fragments from naïve and microbe-incubated plasma samples. These results, together with the EDTA- or protease inhibitor-mediated inhibition of bacteria phagocytosis by hemocytes, suggested the presence in this protostome species of a sophisticated C3-centered opsonic defense system homologous to the deuterostome complement system.

4 Echinodermata

Echinoderms, which belong to the deuterostome lineage, diverged from chordates about 900 mya. The first suggestion that an immune response mediated by complement molecules might be present in this phylum dates back to the pioneering work of Bertheussen and coworkers (Bertheussen 1982 and 1983; Bertheussen and Seljelid 1982). In a series of experiments, they demonstrated that phagocytosis of red blood cells by echinoderm coelomocytes was increased by human C3 and decreased by inhibitors of complement opsonization. These results represented the first evidence suggesting the presence of an alternative complement pathway in an invertebrate species. These functional observations were later confirmed and extended by the identification in a *Strongylocentrotus purpuratus* LPS-activated coelomocyte EST library of two genes (*Sp152* and *Sp064)* encoding SpBf and SpC3, which are homologs, respectively, of the vertebrate complement components Bf/C2 and C3 (Smith et al. 1996).

Sp152 is present in the genome as a single-copy gene that encodes a protein with a deduced molecular mass of 91 kDa and a conserved cleavage site for a putative factor D protease. Analysis of the deduced amino acid sequence revealed that SpBf is

a mosaic protein, including five short consensus repeats (SCR), a VWF domain, and a serine protease domain (Smith et al. 1998). *Sp152* is constitutively expressed in coelomocytes, in which four types of splice variants have been identified by RT-PCR analysis: full-length messages with five SCRs, two kinds of spliced variants with four SCRs (with SCR1 or SCR4 spliced out), and messages with three SCRs (Smith et al. 1998; Smith et al. 2001). No experimental data are yet available with regard to the size and function of SpBf protein/s or their expression pattern in phagocytes. It has been speculated that, by analogy to the mammalian alternative pathway, SpBf interacts with SpC3 to form a C3 convertase.

This hypothesis was supported by the finding in a coelomocytes EST library of a second clone, the aforementioned *Sp064*, *which* encodes SpC3, a molecule exhibiting all the critical molecular and functional features of the vertebrate C3 molecules (Smith et al. 1996; Al-Sharif et al. 1998; Smith et al. 2001). In sea urchins, SpC3 is synthesized as a pro-protein in the coelomocytes. These cells are the only source of C3 in *S. purpuratus*, which lacks organs corresponding to the liver and hepatopancreas. The processed protein, partially purified from coelomic fluid and analyzed by SDS-PAGE, showed a single band of about 210 kDa under non-reducing conditions and two bands of about 130 and 80 kDa under reducing conditions. These bands correspond, respectively, to the typical α and β chains of the vertebrate C3 (Al-Sharif et al. 1998; Smith et al. 2001). The SpC3 α chain has a conserved thioester site (GCGEQ) in the context of a hydrophobic region, as well as a catalytic histidine located about 100 amino acids toward the C terminus. Experiments carried out with methylamine have suggested that SpC3 exhibits functional characteristics that are typical of thioester-mediated opsonic activity, such as methylamine binding and autolysis (Smith 2002).

More recently, the finding that SpC3 represents a major humoral opsonin in *S. purpuratus* coelomic fluid has extended these results. In fact, the phagocytic activity of coelomocytes increases after the incubation of yeast target cells with coelomic fluid containing SpC3, and this activity can be specifically inhibited with an anti-SpC3 antibody (Clow et al. 2004).

Among the four phagocyte types identified in the total coelomocyte population of *S. purpuratus* (Johnson 1969), only a large discoidal type and a smaller polygonal form express SpC3, with different subcellular localization (Gross et al. 2000). An analysis using confocal microscopy revealed that only a single coelomocyte type, the polygonal phagocyte, is able to ingest the opsonized yeast cells (Clow et al. 2004). The expression of SpC3 in the coelomic fluid is increased in response to LPS injection or to injury, with a slightly greater increase in response to LPS. Coelomocytes exhibit similar behavior: in response to LPS injection or to injury, their concentration increases in coelomic fluid, as does the number of SpC3+ coelomocytes (Clow et al. 2000).

Screening efforts to detect *Sp064* transcripts in *S. purpuratus* during embryonic development have revealed the presence of message in unfertilized eggs and throughout embryogenesis, with peak levels at the mesenchyme blastula and gastrula stages. Continuous exposure of the embryos to a heat-killed marine pathogen, *Vibrio diazotrophicus*, from the hatched blastula stage to the pluteus stage, produces a

significant increase in *Sp064* transcripts when compared to the unexposed embryos. These results suggest that in the case of the sea urchin, the developing embryo has a defense system against pathogens that is operated by the complement system (Shah et al. 2003).

In addition to *Sp152* and *Sp064*, two other complement-related cDNAs, *SP5* and *SP5013*, have been identified in *S. purpuratus* coelomocytes. These genes encode two mosaic proteins, SpCRL (*Strongylocentrotus purpuratus* complement related protein, long form) and SpCRS (*Strongylocentrotus purpuratus* complement related protein, short form), which possess structural domains that are also found in the regulatory proteins factor H and factor I and the terminal pathway molecules C6 and C7. The functional roles of these two genes, which are constitutively expressed in all sea urchin tissues, are still unknown (Multerer and Smith 2004).

5 Chordata

The phylum Chordata includes three subphyla: the Cephalochordata, which, according to the most recent phylogenetic analyses (Delsuc et al. 2006; Vienne and Pontarotti 2006), diverged first from a common chordate ancestor around 890 mya; and the Urochordata and Vertebrata, which diverged from each other about 790 mya (Nonaka and Kimura 2006).

Despite the fact that cephalochordates and urochordates occupy a key phylogenetic position in the evolution of immune-related molecules and mechanisms, very little information was available until recently about the function, molecules, and pathways of the complement system in these chordates. While this lack of information still holds true for cephalochordates, the situation has changed significantly for urochordates, in part because of the impetus provided by the sequencing of the *Ciona intestinalis* genome (Dehal et al. 2002) and the ongoing genome projects in *Ciona savignyi* and *Oikopleura dioica*.

5.1 Cephalochordata

The information available regarding the cephalochordate complement system pertains mainly to the identification and sequencing of two mannose-binding lectin-associated serine proteases (MASPs) (MASP-1 and MASP-3; Endo et al. 2003) and of C3-like (amphiC3) and C6-like (amphiC6) cDNA clones from a notochord cDNA library of the amphioxus *Branchiostoma belcheri* (Suzuki et al. 2002).

The amphioxus MASP-1/3 gene structure is very similar to that of human MASP-1/3 (Fujita 2002). It consists of a region of eight exons encoding a heavy (H) chain, followed by a single exon encoding a MASP-3 light (L) chain and a five-exon region encoding a MASP-1 L chain. MASP-1 and MASP-3 are generated as a result of alternative splicing of the primary mRNA, producing two pro-enzymes with identical H chains and distinct L chains, and with the conserved six-domain structure of the human MASPs (Endo et al. 2003). According to the primary structure and exon organization of the genes and the codon encoding the active serine site, the MASP

gene family has been divided into two lineages, the TCN and the AGY types (Endo et al. 1998). The amphioxus MASP-1 and MASP-3 serine protease domains are characterized by an active site encoded by an AGY codon, unusually combined with a disulfide bridge forming a histidine loop (Endo et al. 2003). The combination found in amphioxus is intermediate between the TCN-type, characteristic of the human MASP-1 and ascidian MASPs (Ji et al., 1997) (possessing a TCN codon at the active serine site and a histidine loop disulfide bridge in the serine protease domain) (Arlaud and Gagnon 1981), and the AGY-type, which includes MASP-2, MASP-3, C1r/C1s, and lower vertebrate MASPs (characterized by the absence of a histidine loop disulfide bridge and the presence of an AGY codon at the active serine site). Another peculiar feature of amphioxus MASPs is the presence of an aspartic acid residue at position −6 upstream of the active serine site, which suggests trypsin-type substrate specificities (Kraut 1977).

The *amphiC3* full-length clone in *B. belcheri* encodes a protein of 1792 amino acids, exhibiting 29% identity with human C3 and containing a possible β-α processing site, a canonical thioester site, and the downstream catalytic histidine. A canonical anaphylatoxin domain was also predicted, with six cysteine residues in conserved positions (Suzuki et al. 2002).

Conversely, the full-length *amphiC6* clone encodes a C6-like molecule with the highest identity to human C6. This C6-like molecule has a conserved modular structure in the central portion of the molecule, including two thrombospondin type1 (TSP), one low-density lipoprotein-receptor class A (LDLRA), one MAC/perforin, one epidermal growth factor (EGF), and one TSP module, as well as sequences in the N- and C-terminal regions with no significant similarity to any known sequence. The N- and C-terminal regions also present unique features not found in any molecule of the lytic pathway: arginine and proline stretches at the N-terminus, and nine consecutive repeats of the heptapeptide DA(D/E)TSPG at the C-terminus (Suzuki et al. 2002).

The only functional data available concern the hemolytic activity exhibited by *B. belcheri* humoral fluid toward rabbit erythrocytes; this activity is both Mg^{2+}-dependent and heat-sensitive. The fact that C3 is present in the humoral fluid and that the hemolytic activity can be inhibited by a rabbit anti-human C3 antiserum, zymosan, methylamine, hydrazine, and phenylmethylsulfonyl fluoride provides indirect evidence for the presence of a complement system-mediated immune response (Zhang et al. 2003).

5.2 Urochordata

The urochordate subphylum has been organized into three classes: Larvacea, Thaliacea and Ascidiacea. Species belonging to the Ascidiacea include well-established models that are used by a large community of researchers worldwide. In fact, because of their wide geographical distribution and phylogenetic position, which have allowed researchers to explore the evolutionary origins of the vertebrate lineage, ascidians are among the most extensively studied animal models, especially in the fields of developmental biology and, more recently, immunology.

For studies of the complement system, two main models have been used: the Japanese species *Halocynthia roretzi* and the temperate seawater cosmopolitan species *Ciona intestinalis*. More limited information has also come from two Australian species, *Pyura stolonifera* and *Styela plicata*.

Genomic analyses to detect immune-related genes in *C. intestinalis* (Dehal et al. 2002; Azumi et al. 2003) have provided the most comprehensive picture of complement genes in urochordates, confirming data already reported in the literature and drastically expanding our knowledge at the molecular level. By using a pattern-based search method (Azumi et al. 2002), a number of genes encoding components of the lectin and alternative activation pathways have been identified: nine mannose-binding lectins, nine ficolins, two C1q molecules, four members of the C3/C4/C5/alpha-2-macroglobulin (α2M) family, two factorB/C2 components, and four MASPs. This analysis, extended to the late complement components, has allowed the identification of eleven gene models containing the MAC/perforin domain. Nine of these, exhibiting domain structures similar to those of human late components, are potential complement components. A search of the *Ciona* genome for complement regulatory components, characterized in mammals by SCR domain repeats, has resulted in the identification of 132 gene models (Azumi et al. 2002).

The most relevant result of this analysis concerns the unpredictable molecular complexity of the ascidian complement system, which exhibits an expansion in gene number that is comparable to or, in some cases, even higher than that of its mammalian counterparts. This phylogenetic analysis indicates that gene expansion was generated by duplication events that occurred independently in the ascidian and vertebrate lineages.

Detailed information on individual ascidian complement components at both the molecular and functional levels has come from the studies summarized in the following paragraphs.

5.2.1 C3, C3a, C3aR

C3 genes have been sequenced and characterized in two different urochordate species, the ascidians *H. roretzi* (Nonaka et al. 1999) and *C. intestinalis* (Marino et al. 2002), which contain one (*AsC3*) and two C3-like genes (*CiC3-1* and *CiC3-2*), respectively. The deduced amino acid sequences of *H. roretzi* and *C. intestinalis* C3s exhibit a canonical processing site for α and β chains, as well as a thioester site, with a catalytic histidine located downstream. A convertase cleavage site is present in a conserved position in AsC3 and CiC3-1, while the presence of a long sequence insertion, rich in threonine residues in the same region, makes it difficult to locate the site in CiC3-2.

In mammals, the main site of C3 production is the liver; in ascidians, which lack a true liver, no screening has been done to identify C3-producing tissues. However, Northern blot analysis in *H. roretzi* has revealed C3 transcripts in the hepatopancreas and blood cells (Nonaka et al, 1999), and *CiC3-1* and *CiC3-2* have been cloned from *Ciona* hemocyte RNA (Marino et al, 2002).

The presence of the C3 gene products in the circulating hemolymph of *H. roretzi* and *C. intestinalis* has been established using specific antibodies, which in *Ciona* also identified compartment cells and granular amoebocytes as the C3-producing hemocyte cell types in sections of the tunic of LPS-injected animals (Pinto et al. 2003). Similar results have been obtained in *Styela plicata*, in which heterologous antibodies revealed in the hemolymph a C3-like protein (Raftos et al. 2002) that was synthesized by phagocytic hemocytes and rapidly exocytosed after stimulation with pathogen-associated antigens (Raftos et al. 2004).

The *H. roretzi* C3 acts as a humoral opsonin. In fact, incubation of yeast target cells with hemolymph containing *H. roretzi* C3 leads to a significant increase in phagocytic activity by hemocytes. This activity is completely abolished when the hemolymph is depleted of C3 by incubation with an anti-AsC3 antibody or with a chelating agent such as EDTA. The removal of either C3 or divalent cations from the hemolymph also abolishes the binding of C3 to the yeast target cells, thus demonstrating that phagocytosis proceeds through the yeast opsonization mediated by C3, whose activation requires divalent cations (Nonaka et al. 1999).

In mammals, production of the C3b fragment of the opsonic pathway is associated with the release of the anaphylatoxin C3a, a potent mediator of inflammatory reactions. Mammalian C3a includes six conserved cysteines that form three disulfide bonds which stabilize a tightly packed core consisting of four antiparallel helical regions (Ember et al. 1998; Wetsel et al. 2000). The ascidian C3s sequenced thus far have only four conserved cysteines and do not possess a canonical anaphylatoxin domain. Furthermore, the ascidian C3a C-terminal sequences, like those of all other known invertebrates and lower vertebrates, differ from the mammalian C-terminal consensus sequence GLAR, which participates in binding to the C3a receptor (Ember et al. 1998; Wetsel et al. 2000). Despite these differences in structure, C3a-mediated inflammatory activity has been demonstrated in *C. intestinalis* (Pinto et al. 2003). In fact, both the recombinant C3-1a fragment and synthetic peptides reproducing the C-terminus, with and without the terminal arginine, are able to promote *in vitro* hemocyte chemotaxis in a dose-dependent manner. This activity can be inhibited by an anti-*Ciona* C3-1a-specific antibody, as well as by the pretreatment of hemocytes with pertussis toxin. These results clearly indicate that the chemotactic activity is dependent on the interaction between the ligand C3a and a G protein-coupled receptor.

C. intestinalis C3a activity has been further characterized by immunohistochemical and *in situ* hybridization studies of tunic sections of LPS-injected animals. These analyses showed a total hemocyte number that was 5 times higher in the injured area in LPS-injected animals than in the controls and that the number of granular amoebocytes was increased 15-fold in the injured area. At the same time, granular amoebocytes and compartment cells were actively engaged in producing C3, which reached its highest level of expression at 48 h after LPS injection. Taken together, these findings indicate that C3-1a acts as a chemotaxin for *C. intestinalis* hemocytes and that C3-1a-mediated hemocyte recruitment to sites of injury may play an important role in inflammatory processes (Pinto et al. 2003).

Similar results have been obtained using *Pyura stolonifera*, another species of Ascidiacea. Activation of serum by LPS and zymosan in this species generates an 8.5-kDa proteolytic fragment that confers on serum a chemotactic activity toward hemocytes, as demonstrated by *in vitro* chemotaxis experiments (Raftos et al. 2003).

In mammals, the bioactive fragment C3a binds specifically to cell-surface G protein-coupled seven-transmembrane receptors (Ember et al. 1998; Wetsel et al. 2000). Recently, the receptor molecule involved in C3a-mediated chemotaxis (CiC3aR) has been cloned and characterized in *C. intestinalis* (Melillo et al. 2006). Its expression profile, as evaluated by Northern blot analysis, indicates that like the mammalian C3a receptors, it is broadly expressed in different tissues and organs. It encodes a 95-kDa seven-transmembrane protein that is characterized by a long hydrophilic region of 162 amino acids located between the fourth and fifth transmembrane domains, a common feature of all mammalian C3a receptors. The secondary structure prediction and alignment with other C3aRs evidenced the presence of a very long insertion in the third cytoplasmic loop and an elongation of the cytoplasmic tail. Immunostaining of circulating hemocytes performed with three polyclonal antibodies raised against synthetic peptides reproducing sequences of the first and second extracellular loops and the third intracellular loop have revealed that CiC3aR is constitutively expressed in only two kinds of phagocytic hemocytes, hyaline and granular amoebocytes. In chemotactic assays, the antibodies against the first and second extracellular loops can inhibit the directional migration of hemocytes toward the synthetic peptide reproducing the CiC3a C-terminal sequence, thus providing compelling evidence that *C. intestinalis* expresses a functional C3aR homologous to the mammalian receptor (Melillo et al. 2006).

5.2.2 Factor B

Thus far, factor B (Bf) genes have been recognized, sequenced and analyzed in detail at molecular level only in *C. intestinalis* (Dehal et al. 2002; Azumi et al. 2003; Yoshizaki et al. 2005).

Three genes identified in this species, *CiBf-1*, *CiBf-2*, and *CiBf-3*, encode proteins with identical domain structures that resemble the basic domain structure of the vertebrate Bf/C2 gene family. They are characterized by the presence of three SCR domains, a von Willebrand factor type A domain, and a serine protease domain. In addition to these, in *Ciona* two LDLR domains and one SCR domain are present at the N-terminus. The active site of the three CiBf serine protease domains is of the AGY type, a feature shared with vertebrate MASP-2, MASP-3, and C1r/C1s, together with the absence of a histidine loop disulfide bridge. The amino acid sequence identity between CiBf-1 and CiBf-2 is 88%, and between CiBf-3 and CiBf-1 or CiBf-2 is 49%.

Phylogenetic analysis, including a detailed characterization of the genomic organization and intron/exon composition, has indicated that CiBf genes are the result of duplication and gene conversion events that occurred within the urochordate lineage after the divergence from the vertebrate subphylum. CiBf genes, distributed within a 50-kb genomic region, mapped to a different chromosome than do the

CiC3-1 and CiC3-2 loci, suggesting that the linkage among the MHC class III complement genes was established in vertebrates.

5.2.3 MBL/Ficolins

Clear evidence of the presence of a lectin activation pathway in ascidians has been provided by the molecular and functional characterization of an *H. roretzi* 36-kDa lectin (GBL) that binds specifically to glucose. Sequence analysis revealed that GBL contains a C-type lectin domain but lacks the collagen-like domain of mammalian MBLs.

To determine whether GBL acts as a recognition molecule in the lectin pathway of the ascidian complement system, the *Halocynthia* plasma was depleted of GBL or C3 with specific antibodies and used in assays of yeast phagocytosis by hemocytes. Incubation of plasma with either of the antibodies resulted in a significant decrease in the phagocytic activity, suggesting that GBL recognizes carbohydrates on the yeast surface and in turn activates C3 through the associated MASPs (Sekine et al. 2001).

In mammals, the lectin pathway can be triggered by pathogens and involves recognition by lectins other than MBL. These lectins belong to the ficolin family and are characterized by the presence of an NH_2-terminal domain containing cysteine residues, a fibrinogen-like domain, and a collagen-like domain (Matsushita et al. 2000). In *H. roretzi,* four cDNA clones encoding orthologs of mammalian ficolins have been sequenced and their carbohydrate binding specificity has been assessed. It has been suggested that their association with MASPs contributes to the activation of C3, but thus far there is no experimental evidence to support this hypothesis (Kenjo et al. 2001).

5.2.4 MASPs

Two MASPs have been identified in an *H. roretzi* cDNA library, MASPa and MASPb. The deduced amino acid sequences of these two molecules exhibit the two main specific features of mammalian MASP1: the codon of the serine in the active site is of the TCN type, and the catalytic histidine is present in a disulfide bridge loop. The residue determining the substrate specificity is located at the -6 position with reference to the active site; this residue is aspartic acid in MASPa, as in all vertebrate MASPs, and threonine in MASPb. This finding suggests a trypsin-like activity for MASPa and a different substrate specificity for MASPb. The absence of experimental data clarifying the activity of the two ascidian MASPs means that it is not currently possible to interpret the sequence analysis results.

5.2.5 Complement Receptors Type 3 and Type 4

In mammals, complement receptors type 3 and type 4 (CR3 and CR4) are members of the β_2-integrin family. These molecules are membrane-bound heterodimers, each consisting of a different α subunit non-covalently associated with the same β subunit.

Among the apparently unrelated functions documented for CR3 and CR4 is the ability of both molecules, expressed mainly on myeloid cells, to stimulate phagocytosis by binding iC3b-opsonized bacteria.

In *H. roretzi*, four integrin subunits, two α (α_{Hr1} and α_{Hr2}) and two β (β_{Hr1} and β_{Hr2}), have been cloned, sequenced and partially characterized (Miyazawa et al. 2001; Miyazawa and Nonaka 2004). Analysis of the deduced amino acid sequences indicated that all of them have the typical domain structures of the mammalian α and β integrin subunits. Involvement of the integrin α_{Hr1} subunit in the C3-dependent phagocytic activity of *H. roretzi* hemocytes has been demonstrated in phagocytosis assays using a specific antibody against a recombinant protein reproducing the extracellular region of α_{Hr1} (Miyazawa et al. 2001).

To investigate the heterodimer composition in ascidians, insect cell lines were co-infected with two recombinant baculovirus species, the first containing the α_{Hr1} gene and the second containing either the β_{Hr1} or β_{Hr2} gene. Immunoprecipitation of the cell line extracts with anti-α_{Hr1} antibody demonstrated that both the β_{Hr1}, and β_{Hr2} subunits are associated with the α_{Hr1} subunit. The association of α_{Hr1} with β_{Hr1} was also confirmed in western blot analysis of ascidian hemocytes.

The type of pairing found in ascidians, namely the same integrin α subunit (α_{Hr1}) paired with different integrin β subunits (β_{Hr1} and β_{Hr2}), is different from the mammalian CR3 and CR4 pairing pattern and resembles that of α_v integrins. Because of this type of pairing, the authors referred to these ancestral forms of complement receptors as the "α_{Hr1} integrin family" or "hemocyte integrin family."

6 Concluding Remarks

The search for complement components in the invertebrate species analyzed thus far has clearly demonstrated the presence of C3-like and Bf-like molecules in very ancient animal phyla and has identified these molecules as the most basic complement component assembly. In particular, comparative analysis of C3 amino acid sequences has identified the most important molecular signatures in very primitive species. In fact, the thioester site with its associated catalytic histidine, the β-α chain processing site, the C3 convertase site, and the anaphylatoxin domain, are features shared by all C3 molecules. These findings strongly indicate the recruitment, very early in evolution, of ancestral C3 and Bf into a primordial alternative pathway of C3 activation, with the consequent formation of the peptide fragment C3a and the opsonin fragment C3b, which binds to the microbial surface through a thioester bond. In this context, it is remarkable that the human C3a fragment has antimicrobial properties. It binds to bacterial membranes and directly kills the microorganisms by inducing breaks in the membrane. These results suggest an unforeseen ancestral role for a C3-related peptide and point to the C3a molecule as a link between two important arms of innate immunity, the complement system and antimicrobial peptides (Andersson Nordahl et al. 2004). Among the many different roles played by

the C3a peptide in mammals (Mastellos and Lambris 2002), antimicrobial activity could represent the simplest and most primordial function. Studies of invertebrate complement components, which lack the complexity of the mammalian complement system, could assist us in exploring this activity and discovering new molecular mechanisms underlying the central event of the complement reaction cascade, namely C3 activation. The study of invertebrates can further improve our knowledge of the complement system, which is surprisingly flexible, as shown by the very recent finding in mouse of a novel C3 activation pathway mediated by SIGN-R1, a membrane-bound lectin, which initiates a classical but Ig-independent pathway (Kang et al. 2006). The analysis of the urochordate *C. intestinalis* genome has confirmed the absence of the pivotal genes for adaptive immunity in this species, which is phylogenetically at the base of the vertebrate lineage. At the same time, it has confirmed the presence of many complement genes, including two C1q-like genes. The ascidian C1q could either act as a lectin, like the lamprey C1q (Matsushita et al. 2004), or with other unknown membrane bound receptor/s, as in the case of murine SIGN-R1. Future investigation of ascidian C1q expression and biological function could help to shed light on the origin of the classical pathway of the complement system.

It is noteworthy that the studies carried out on the invertebrate complement system are still fragmentary, limited to a few species and available genomes, and mainly devoted to genomic and cDNA sequence analyses. Greater efforts in investigating the biological activities of invertebrate complement components and their mutual functional relationships should provide answers to the many basic questions in this field that remain to be answered.

7 Acknowledgements

References

Al-Sharif, W.Z., Sunyer, J.O., Lambris, J.D. and Smith, L.C. (1998) Sea urchin coelomocytes specifically express a homologue of the complement component C3. J. Immunol. 160, 2983-2997.

Andersson Nordahl, E., Rydengård, V., Nyberg, P., Nitsche, D.P., Mörgelin, M., Malmsten, M., Björck, L. and Schmidtchen, A. 2004 Activation of the complement system generates antibacterial peptides. Proc. Natl. Acad. Sci. USA 101, 16879-16884.

Azumi, K., De Santis, R., De Tomaso, A., Rigoutsos, I., Yoshizaki, F., Pinto, M.R., Marino, R., Shida, K., Ikeda, M., Ikeda, M., Arai, M., Inoue, Y., Shimizu, T., Satoh, N., Rokhsar, D.S., Du Pasquier, L., Kasahara, M., Satake, M., and Nonaka, M. (2003). Genomic analysis of immunity in a Urochordate and the emergence of the vertebrate immune system: "waiting for Godot". Immunogenetics 55, 570-581.

Bertheussen, K. (1982) Receptors for complement on echinoid phagocytes. II. Purified human complement mediates echinoid phagocytosis. Dev. Comp. Immunol. 6, 635-642.

Bertheussen, K. and Seljelid R. (1982) Receptors for complement on echinoid phagocytes. I. The opsonic effect of vertebrate sera on echinoid phagocytosis. Dev. Comp. Immunol. 6, 423-431.

Bertheussen, K. (1983) Complement-like activity in sea urchin coelomic fluid. Dev. Comp. Immunol. 7, 21-31.

Clow, L.A., Gross, P.S., Shih, C-S. and 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. and Smith, L.C. (2004) The sea urchin complement homologue, SpC3, functions as an opsonin. J. Exp. Biol. 207, 2147-2155.

Dehal P., Satou, Y., Campbell, R.K., Chapman, J., Degnan, B., De Tomaso, A. Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein D.M. et al. 2002. The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. Science 298, 2157-2167.

Delsuc, F., Brinkmann, H., Chourrout, D. and Philippe, H. (2006) Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature 439, 965-968.Edds, K. (1993) Cell biology of echinoid coelomocytes. I Diversity and characterization of cell types. J. Invert. Pathol. 61, 173-178.

Dishaw, L.J., Smith, S.L. and Bigger, C.H. 2005 Characterization of a C3-like cDNA in a coral: phylogenetic implications. Immunogenetics 57, 535-548.

Ember, J.A., Jagels, M.A. and Hugli, T.E. (1998). Characterization of complement anaphylatoxins and their biological responses. In: J.E. Volanakis and M.M. Frank, (Eds.) *The Human Complement System in Health and Disease*. Marcel Dekker, New York, pp. 241-284.

Endo, Y., Nonaka, M., Saiga, H., Kakinuma, Y., Matsushita, A., Takahashi, M., Matsushita, M. and Fujita, T. (2003) Origin of mannose-binding lectin-associated serine protease (MASP)-1 and MASP-3 involved in the lectin complement pathway traced back to the invertebrate, amphioxus. J Immunol. 170, 4701-4707.

Fujita, T. (2002) Evolution of the lectin-complement pathway and its role in innate immunity. Nat. Rev. Immunol. 2, 346-353.

Gross, P.S., Clow, L.A. and 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.

Ji, X, Azumi, K, Sasaki, M. and Nonaka M. (1997) Ancient origin of the complement lectin pathway revealed by molecular cloning of mannan binding protein-associated serine protease from a urochordate, the Japanese ascidian, *Halocynthia roretzi*. Proc. Natl. Acad. Sci. USA 94, 6340-6345.

Johnson, P.T. (1969) The coelomic elements of sea urchins (*Strongylocentrotus*). I. The normal coelomocytes; their morphology and dynamics in hanging drops. J. Invert. Pathol. 13, 25-41.

Kairies, N., Beisel, H.G., Fuentes-Prior, P., Tsuda, R., Muta, T., Iwanaga, S., Bode, W., Huber, R. and Kawabata, S. (2001) The 2.0-A crystal structure of tachylectin 5A provides evidence for the common origin of the innate immunity and the blood coagulation systems. Proc. Natl. Acad. Sci. U S A 98, 13519-13524.

Kang, Y-S., Do, Y., Lee, H-K., Park, S.H., Cheong, C., Lynch, R.M., Loeffler, J.M., Steinman, R.M. and Park, C.G. 2006 A dominant complement fixation pathway for pneumococcal polysaccharides initiated by SIGN-R1 interacting with C1q. Cell 125, 47-58.

Karp, D.R., Parker, K.L., Shreffler, D.C. and Capra, J.D. (1981) Characterization of the murine C4 precursor (pro-C4): evidence that the carboxyterminal subunit is the C4 gamma-chain. J. Immunol. 126, 2060-2061.

Kelly, K.L., Cooper, E.L. and Raftos, D.A. (1992) Purification and characterization of a humoral opsonin from the solitary urochordate *Styela clava*. Comp. Biochem. Physiol. 103B, 749-753.

Kenjo, a., Takahashi, M., Matsushita, M., Endo, Y., Nkata, M., Mizuochi, T. and Fujita, T. (2001) Cloning and characterization of novel ficolins from the solitary ascidian, Halocynthia roretzi. J. Biol. Chem.276, 19959-19965.

Kraut, J. (1977) Serine proteases: structure and mechanism of catalysis. Annu. Rev. Biochem. 46, 331-358.

Lambris, J.D. (1990) The Third Component of Complement - Chemistry and Biology. Springer-Verlag, Heidelberg.

Marino R., Kimura, Y., De Santis, R., Lambris, J.D. and Pinto M.R. (2002). Complement in urochordates: cloning and characterization of two C3-like genes in the ascidian *Ciona intestinalis*. Immunogenetics 53, 1055-1064.

Mastellos, D. and Lambris, J.D. (2002) Complement: more than a 'guard' against invading pathogens? Trends Immunol. 23, 485-91.

Matsushita, M., Endo, Y., Nonaka, M. and Fujita, T. (1998) Complement-related serine proteases in tunicates and vertebrates. Curr. Opin. Immunol. 10, 29-35.

Matsushita, M., Endo, Y. and Fujita, T. (2000) Cutting edge: complement-activating complex of ficolin and mannose-binding lectin-associated serine protease. J. Immunol. 164, 2281-2284.

Matsushita, M. and Fujita, T. (2002) The role of ficolins in innate immunity. Immunobiology 205, 490-497.

Matsushita, M., Matsushita, A., Endo, Y., Nakata, M., Kojima, N., Mizuochi, T. and Fujita, T. (2004) Origin of the classical complement pathway: lamprey orthologue of mammalian C1q acts as a lectin. Proc. Natl. Acad. Sci. USA 101, 10127-10131.

Melillo, D., Sfyroera, G., De Santis, R., Graziano, R., Marino, R., Lambris, J.D. and Pinto, M.R. (2006) First identification of a chemotactic receptor in an invertebrate species: structural and functional characterization of *Ciona intestinalis* C3a Receptor. J. Immunol. 177, 4132-4140.

Miyazawa, S., Azumi, K., and Nonaka, M. (2001) Cloning and characterization of integrin α subunits from the solitary ascidian, *Halocynthia roretzi*. J. Immunol. 166, 1710-1715.

Miyazawa, S. and Nonaka, M. (2004) Characterization of novel ascidian β integrins as primitive complement receptor subunits. Immunogenetics 55, 836-844.

Multerer, K.A. and Smith, L.C. (2004) Two cDNAs from the purple sea urchin, Strongylocentrotus purpuratus, encoding mosaic proteins with domains found in factor H, factorI, and complement components C6 and C7. Immunogenetics 56, 89-106.

Nonaka, M. (1994) Molecular analysis of the lamprey complement system. Fish Shellfish Immunol. 4, 437-446.

Nonaka, M., Azumi, K., Ji, X., Namikawa-Yamada, C., Sasaki, M., Saiga, H., Dodds, A.W., Sekine, H., Homma, M.K., Matsushita, M., Endo, Y. and Fujita, T. (1999) Opsonic complement component C3 in the solitary ascidian, *Halocynthia roretzi*. J. Immunol. 162, 387-391.

Nonaka, M. and Kimura, A. (2006) Genomic view of the evolution of the complement system. Immunogenetics 58, 701-713.

Pearce, S., Newton, R.A., Nair, S.V. and Raftos, D.A. (2001) Humoral opsonins of the tunicate, *Pyura stolonifera*. Dev. Comp. Immunol. 25, 377-385.

Pinto M.R., Chinnici, C.M., Kimura, Y., Melillo, D., Marino, R., Spruce, L.A., De Santis, R., Parrinello, N. and Lambris, J. D. (2003) CiC3-1a-mediated chemotaxis in the deuterostome invertebrate *Ciona intestinalis* (Urochordata). J Immunol. 171, 5521-5528.

Raftos, D.A., Nair, S. V., Robbins, J., Newton, R. A. and Peters. R. (2002) A complement component C3-like protein from the tunicate, *Styela plicata*. Dev. Comp. Immunol. 26, 307-312.

Raftos, D.A., Robbins, J. Newton, R.A. and Nair, S. V. (2003) A complement component C3a-like peptide stimulates chemotaxis by hemocytes from an invertebrate chordate-the tunicate, *Pyura stolonifera*. Comp. Biochem. Physiol. Part A 134, 377-386.

Raftos D.A., Stillman, D.L. and Cooper, E. L. (1998) Chemotactic responses of tunicate (Urochordata, Ascidiacea) hemocytes in vitro. J. Invertebr. Pathol. 72**,** 44-49.

Sekine, H., Kenjo, A., Azumi, K., Ohi, G., Takahashi, M., Kasukawa, R., Ichikawa, N., Nakata, M., Mizuochi, T., Matsushita M., Endo, Y. and Fujita, T. (2001) An ancient lectin-dependent complement system in an ascidian: novel lectin isolated from the plasma of the solitary ascidian, *Halocynthia roretzi*. J. Immunol. 167, 4504-4510.

Shah, M., Brown, K.M. and 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.

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., Britten, R.J. and Davidson E.H. (1995) Lipopolysaccharide activates the sea urchin immune system. Dev. Comp. Immunol. 19, 217-224.

Smith, L.C., Chang, L., Britten, R.J. and Davidson, E.H. (1996) Sea urchin genes expressed in activated coelomocytes are identified by expressed sequence tags. J. Immunol. 156, 593-602.

Smith, L.C., Clow, L.A. and Terwilliger, D.P. (2001) The ancestral complement system in sea urchins. Immunol. Rev. 180, 16-34.

Smith, L.C., Shih, C-S. and 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.

Song, W-C., Sarrias, M.R. and Lambris, J.D. (2000) Complement and innate immunity. Immunopharmacology 49, 187-198.

Suzuki, M.M., Satoh, N. and 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.

Vienne, A. and Pontarotti, P. (2006) Metaphylogeny of 82 gene families sheds a new light on chordate evolution.Int. J. Biol. Sci. 2, 32-37.

Vogel, C.W., Bredehorst, R., Fritzinger, D.C., Grunwald, T., Ziegelmuller, P. and Kock, M.A. (1996) Structure and function of cobra venom factor, the complement-activating protein in cobra venom. Adv. Exp. Med. Biol. 391, 97-114.

Wetsel, R.A., Kildsgaard, J. and Haviland, D.L. (2000) Complement anaphylatoxins (C3a, C4a, C5a) and their receptors (C3aR, C5aR/CD88) as therapeutic targets in inflammation. In: J.D. Lambris and V.M. Holers, (Eds.), *Therapeutics Interventions in the Complement System*. The Humana Press, New York, pp. 113-153.

Yoshizaki, F.Y., Ikawa, S., Satake, M, Satoh, N., Nonaka, M. (2005) Structure and the evolutionary implication of the triplicated complement factor B genes of a urochordate ascidian, *Ciona intestinalis*. Immunogenetics 56, 930-942.

Zhu, Y., Thangamani, S., Ho, B. and Ding, J.L. (2005) The ancient origin of the complement system. EMBO J. 24, 382-394.