Rita Marino · Yuko Kimura · Rosaria De Santis
John D. Lambris · Maria Rosaria Pinto

Complement in urochordates: cloning and characterization of two *C3*-like genes in the ascidian *Ciona intestinalis*

Abstract The recent identification of complement components in deuterostome invertebrates has indicated the presence of a complement system operating via an alternative pathway in echinoderms and tunicates and via a MBL-mediated pathway thus far identified only in tunicates. Here, we report the isolation of two *C3*-like genes, *CiC3-1* and *CiC3-2*, from blood cell total RNA of the ascidian *Ciona intestinalis*. The deduced amino acid sequences of both *Ciona* C3-like proteins exhibit a canonical processing site for α and β chains, a thioester site with an associated catalytic histidine and a convertase cleavage site, thus showing an overall similarity to the other C3 molecules already characterized. Southern blotting analysis indicated that each gene is present as a single copy per haploid genome. In situ hybridization experiments showed that both *CiC3-1* and *CiC3-2* are expressed in one type of blood cell, the compartment cells. Two polyclonal antibodies, raised against two deduced peptide sequences in the α chain of CiC3-1 and CiC3-2, allowed the identification by Western blot of a single band in the blood serum, of about M_r 150,000. A phylogenetic tree, based on the alignment of CiC3-1 and CiC3-2 with molecules of the α_2-macroglobulin superfamily, indicated that the *Ciona* C3s form a cluster with *Halocynthia roretzi* C3. The phylogenetic analysis also suggested that the duplication event from which the *CiC3-1* and *CiC3-2* genes originated occurred in the urochordate lineage after the separation of the *Halocynthia* and *Ciona* ancestor.

Keywords Ascidians · Complement evolution · C3 · Thioester

R. Marino · R. De Santis · M.R. Pinto (✉)
Laboratory of Cell Biology, Stazione Zoologica "Anton Dohrn",
Villa Comunale, 80121 Naples, Italy
e-mail: pinto@alpha.szn.it
Tel.: +39-081-5833284, Fax: +39-081-7641355

Y. Kimura · J.D. Lambris
Department of Pathology and Laboratory Medicine,
University of Pennsylvania, 401 Stellar Chance, Philadelphia,
PA 19104, USA

Introduction

The ability of the host to defend itself against invading pathogens is universal among metazoans and reaches the highest level of complexity in higher vertebrates, where the integrated action of the innate and adaptive immune systems provides an efficient and prompt immune response. 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). Its effector function can be activated through three different activation pathways: (1) the alternative pathway, which is triggered by the direct binding of a complement component to the pathogen surface; (2) the mannose-binding lectin (MBL) pathway, which is initiated by the binding of a serum lectin to the carbohydrates present on the invading pathogen; and (3) the classical pathway, which is triggered by antibody–antigen binding. In all the complement activation pathways, the third component of the complement, C3, plays a central role (Lambris 1990). Its proteolytic activation by C3 convertase triggers the effector function of complement, leading to the recruitment of inflammatory cells and opsonization of the pathogens or to the lysis of pathogens through the formation of the membrane attack complex (Lambris 1990).

Recently, a C3 homologue, *SpC3* (Al-Sharif et al. 1998), and a factor B homologue, *SpBf* (Smith et al. 1998), were cloned and sequenced from a lipopolysaccharide (LPS)-activated celomocyte cDNA library of the sea urchin *Strongylocentrotus purpuratus* (Echinodermata). Western blot and gene expression analysis by RT-PCR revealed that two *S. purpuratus* phagocyte subpopulations express *SpC3* (Gross et al. 2000), which increases in quantity following LPS challenge (Clow et al. 2000).

Furthermore, studies carried out on the urochordate *Halocynthia roretzi* (Tunicata) revealed the presence of a complement system that comprises at least C3 (Nonaka et al. 1999), two mannan-binding protein-associated serine proteases (Ji et al. 1997), factor B (Nonaka 2001)

and a complement receptor type 3/4 (Miyazawa et al. 2001). Functional studies of *H. roretzi* demonstrated both that its body fluid contains an opsonic activity which enhances the phagocytosis of yeast by ascidian blood cells and that the opsonic activity is inhibited by an antibody against *H. roretzi* C3 (Nonaka et al. 1999). Recently, cDNAs encoding C3 and C6 were cloned and sequenced from a notochord expressed-sequence tag library of the cephalochordate *Amphioxus*, suggesting the presence of the lytic pathway in this much more highly evolved protochordate subphylum (Nonaka 2001).

Taken together, these findings indicate the presence of a simple complement system in deuterostome invertebrates that proceeds via an alternative pathway in echinoderms and tunicates and via a MBL-mediated pathway identified thus far only in tunicates. This system is composed of few elements and shares very strong similarity with a primitive complement cascade (Al-Sharif et al. 1998; Nonaka 2001; Nonaka et al. 1999). These findings also indicate that the complement system emerged at least 600–700 million years ago, long before the appearance of the immunoglobulins, thus pointing to complement as an evolutionary link between the vertebrate and invertebrate immune systems.

The complement system of higher vertebrates consists of more than 30 humoral and cell membrane proteins, some of which have significant sequence similarities. It has been suggested that they originated from a few ancestral molecules either by individual gene duplication events (Bentley 1988) or by extensive genome duplication events that duplicated complement pathways (Kasahara 1999). In particular, it has been suggested that C3, together with C4 and C5, originated from α_2-macroglobulin, a serum protease inhibitor (Sottrup-Jensen et al. 1985). Thus far, only a single C3/C4/C5-like gene has been isolated in the lamprey (Nonaka and Takahashi 1992), the ascidian *H. roretzi* (Nonaka et al. 1999) and the sea urchin *S. purpuratus* (Al-Sharif et al. 1998). These observations led to the conclusion that the gene duplication events that gave rise to the three different homologous genes, *C3*, *C4* and *C5*, occurred later, in the jawed vertebrates.

The ascidian *Ciona intestinalis*, which belongs to the subphylum Urochordata, is a cosmopolitan species closely related to the vertebrate lineage. It provides a good system for studying the evolutionary origins of the chordate lineage, from which vertebrates developed. Its genetic relationship to other invertebrates and vertebrates prompted us to undertake a study of complement in this animal as a mean of better understanding the evolution of *C3*, *C4* and *C5*.

Here, we report the isolation of two C3-like cDNAs, *CiC3-1* and *CiC3-2*, and their pattern of expression in *Ciona* blood cells. Western blot analysis carried out using antibodies raised against the two peptides, synthesized on the basis of the deduced amino acid sequences of *CiC3-1* and *CiC3-2*, confirmed the presence of both gene products in *Ciona* serum.

Materials and methods

RNA isolation

Blood from *C. intestinalis* was collected by syringe in the presence of ~10 mM EDTA, pH 8.0, to prevent cell clotting. After centrifugation for 20 min at 3,000 rpm, the cell pellet was immediately frozen in liquid nitrogen and stored at –80 °C. Total RNA was extracted with the Promega kit SV total RNA isolation system (Promega).

RT-PCR amplification of the thioester-containing region

Random-primed single-strand cDNA was synthesized from blood RNA with an RT-PCR kit (Gibco BRL) and PCR-amplified, using the same degenerate primers (sense primer: 5′-GGNTGYGGN-GARCARAAYATG-3′, antisense primer: 5′-ACRAANGCNGT-NAGCCANGT-3′) and conditions as described for the amplification of lamprey *C3* (Nonaka and Takahashi 1992). A single band of the expected size (230 bp) was detected. This product was cloned into the pCRII-TOPO vector (Invitrogen) and sequenced.

5′- and 3′- rapid amplification of cDNA ends

To clone the sequences at the 5′ and 3′ ends of the two *CiC3*-like genes, RACE assays were carried out using the 5′-RACE and 3′-RACE systems (Gibco BRL). To determine the 3′ ends, two pairs of sense gene-specific primers per gene were designed, according to the sequences of the 230-bp products. To cover the sequences of the two *CiC3*-like genes at the 5′ ends, four cycles of 5′-RACE per gene were performed, using antisense primers designed for the first cycle from the sequences of the 230-bp products and for the following cycles from the sequence of the previous 5′-RACE product.

DNA sequencing and sequence analysis

Sequences were determined on both strands by the dideoxy chain-termination method (Sanger and Coulson 1975), using vector- or gene-specific primers. cDNA sequence analysis was routinely performed by running a basic BLAST search of the GenBank database and the deduced amino acid sequences were aligned using the Clustal X ver 1.8.

The leader cleavage site for CiC3-1 was predicted by the SignalP ver 1.1 software (Nielsen et al. 1997).

To construct a phylogenetic tree, the deduced amino acid sequences of CiC3-1 and CiC3-2, as well as several C3, C4, C5 and α_2-macroglobulin sequences from other species, were aligned using the Clustal W ver 1.5 program (Thompson et al. 1994) and the resulting alignments were manually corrected. The obtained alignment was used to calculate Poisson-corrected distance matrixes to construct trees by the neighbor-joining method (Saitou and Nei 1987) using the MEGA ver 2 program.

Southern blot analysis

Genomic DNA was isolated from *Ciona* sperm according to standard protocols (Sambrook and Russell 2001). The DNA (8 μg) was digested to completion with the restriction enzymes *Eco*RI, *Hin*dIII, *Kpn*I and *Pst*I, separated on a 0.8% agarose gel and transferred to a Hybond-N+ nylon membrane (Amersham Pharmacia Biotech) in 20× saline-sodium citrate (SSC). The transferred DNA was UV-crosslinked at 120 mJ/cm2, pre-hybridized and then hybridized with 32P-labeled random-primed probes (Feinberg and Vogelstein 1983) in a buffer containing 50% formamide, 5× Denhardt's solution, 2× SSC, 50 mM sodium phosphate (pH 6.8), 0.5 mg tRNA/ml and 0.1% SDS at 65 °C. The probes corresponded to the nucleotide fragments 3,955–4,781 for *CiC3-1* (aa 1,311–1,586) and 3,872–4,576 for *CiC3-2* (aa 1,292–1,525).

Following hybridization, the filters were washed extensively at 65 °C in 0.2× SSC and 0.1% SDS and then exposed to X-OMAT AR film (Eastman Kodak, Rochester, N.Y.) for 3 h at –80 °C.

In situ hybridization

Samples of *Ciona* stomach were dissected and fixed in 4% paraformaldehyde in 75% sea water overnight at 4 °C. Samples were dehydrated, embedded in paraffin, sectioned at 7.5 µm and processed for in situ hybridization (Simeone et al. 1995). In situ hybridization was carried out with digoxigenin-11-UTP-labeled riboprobes (Roche Diagnostics), according to the instructions of the manufacturer. The antisense cRNAs corresponded to the nucleotide fragments 3,955–4,781 for *CiC3-1* (aa 1,311–1,586) and 3,872–4,576 for *CiC3-2* (aa 1,292–1,525); 100 ng probe/slide were used at 60 °C. Control experiments were run in parallel using the corresponding sense cRNAs.

Peptide synthesis and antibody production

A 22-aa peptide, CiC3-1$^{936–957}$ (IEDEGELKDIYETFPIDLKNSR; boxed in Fig. 1), representing part of the CiC3-1 α-chain deduced amino acid sequence, and a 21-aa peptide, CiC3-2$^{579–599}$ (RNKRNVELEDSVLLTSLNRKL; boxed in Fig. 1), representing the N-terminal region of the α chain of the deduced amino acid sequence of CiC3-2, were synthesized using an Applied Biosystems 430 A peptide synthesizer (Foster City, Calif.), as previously described (Briand et al. 1985). The peptides were then coupled to ovalbumin by the glutaraldehyde method and used for the immunization of rabbits and the production of antibodies. The anti-peptide-specific antibody was purified by affinity chromatography, using the synthetic peptide coupled to cyanogen bromide-activated Sepharose 4B (Amersham Pharmacia Biotech, Piscataway, N.J.).

Western blot analysis

Blood from *C. intestinalis* was collected by syringe in the presence of ~10 mM EDTA, pH 8.0, and 100 µM phenylmethylsulfonyl fluoride. After centrifugation for 20 min at 3,000 rpm, the supernatant (hereafter referred to as blood serum) was collected and stored at –80 °C.

The reactivity of *Ciona* serum proteins with the anti-CiC3 antibodies was assessed by Western blotting, as previously described (Sarrias et al. 2001). For the detection of the CiC3-1 and CiC3-2 proteins, serum samples (20 µl for the detection of CiC3-1, 40 µl for CiC3-2) were loaded onto 6% SDS-PAGE gels under reducing conditions. The proteins were then electro-blotted onto a polyvinylidene difluoride membrane and the CiC3s were detected using affinity-purified anti-CiC3-1 and anti-CiC3-2 antibodies. Non-specific binding to the membrane was prevented by incubation in blocking buffer (PBS containing 10% milk). Bound polyclonal antibodies were detected with a peroxidase-labeled goat anti-rabbit antibody (Bio-Rad). All of these incubation steps were performed for 1 h at room temperature and the membranes were washed with PBS containing Tween 20 (0.05%) between each incubation step. The proteins were detected using the ECL kit (Amersham Pharmacia Biotech, Piscataway, N.J.).

Accession numbers

Nucleotide sequence data reported here are available in the GenBank database under the accession numbers AJ320542 (*CiC3-1*) and AJ320543 (*CiC3-2*).

Results

Isolation and sequencing of the two *C3*-like genes in *C. intestinalis*

The strategy used to isolate and determine the entire nucleotide sequence of the *Ciona C3* gene included the amplification of the CiC3 thioester region cDNA by RT-PCR of blood total mRNA and 3′-RACE and 5′-RACE analyses.

Random-primed, single-stranded cDNA, synthesized from blood total RNA, was PCR-amplified using two degenerate primers designed according to previous reports (Nonaka and Takahashi 1992) on the basis of two highly conserved amino acid sequences at the thioester site and about 60 aa residues downstream of the thioester site of C3 from various species. The amplification product, a single band of about 230 bp, was sub-cloned into the TOPO-TA-cloning vector. Nucleotide sequence analysis allowed the identification of five different clones, one showing a close similarity to lamprey α$_2$-macroglobulin (51% identity, 69% similarity) and four with close similarity to known C3 molecules. The deduced amino acid sequences of these four *C3*-like clones showed that they encode two different C3-like molecules, termed CiC3-1 and CiC3-2, whose highest percentage of sequence similarity was to sea urchin C3 (41% identity, 51% similarity for CiC3-1; 37% identity, 49% similarity for CiC3-2).

3′-RACE of total RNA from *Ciona* blood was performed to cover gene sequences from the thioester region to the poly(A) tail of the two genes and four different cycles of 5′-RACE were carried out for each gene, in order to cover the sequences from the thioester site to the putative initiation codon of *CiC3-1* and about 90 aa downstream of the initiation codon of *CiC3-2*.

Sequence analyses

The *CiC3-1* sequence we obtained is 5,521 bp in length and encodes a protein of 1,773 aa in its open reading frame. *CiC3-2*, which was partially sequenced, is 5,528 bp in length and encodes a protein of 1,804 aa in the open reading frame. The deduced amino acid sequences of CiC3-1 and CiC3-2 are shown in Fig. 1. Amino acid sequence analyses indicated that the sequence identity/similarity of CiC3-1 to CiC3-2 is 40.7%/51.5%. A potential cleavage site for the post-translational processing of CiC3-1 and CiC3-2 into the α and β chains is present at Arg^{670} of CiC3-1 and Arg^{578} of CiC3-2 (Fig. 1). The β chains have 661 aa residues in CiC3-1 (without leader) and 580 aa residues in CiC3-2 (incomplete at the N-terminus). After removing the RKKR and RNKR at the α-β junctions, the α chains have 999 aa and 1,222 aa residues, respectively.

A long sequence insertion, very rich in Thr residues, was found in the α chain of CiC3-2 (at about 120 aa residues from the α-β cleavage site; Fig. 1). The analysis of

Leader

```
Ciona C3-1      --------MVWFSFSLLVTLLAVATAFDHTVVVPKALRVDALEKIIVNLHGYN-RATITGYLQDLPGLQTFFSRTGQRVLTPAQCQNPIEMTFRVTRDPQGADGIASFGLTQKVRLTIQV  111
Ciona C3-2      -----------------------------------------------------------ISAQQSNTISI---Y---------IVSKNIGRKVEVVISC  28
Halocynthia C3  ----MISFHIYICILFFLTRAKLTVSCSYSLVLPKALRVDTENYAYIDFHNCPPIAKISTSLHSLPGIENTHSSDEK---VDHPTIDPIKLTYVPKSKMMENRKKTTTPYKKALVVINYC
Amphioxus C3    ------MIRSLVCVCLLFAAVLGQQGPAYFISVPNLLKVGTEETVSVNVFNVANPVRVKVYLQDYPDRKTTFSEAEVDVNQDEPSLVTVRVN---P-----DNLPESRATKRYVYVVAKS
Sea urchin C3   MGSIILIFVLAISASLVPGGVTQDPPATFFVISPNVFRAGVEEKVVVTLIRSSPPNVQNIEVKVSLMRPGSTVTFSEDTRLVSPGSSQSFSVVIQAG------DLMRNEGAFQHMVLKAES
Hagfish C3      -----------------VLVIAPAATSSYDDLAVAILMVDQKKITEVHVLLVN-----------PHTGATLDEKKVKLQWDNKFIAFTKLQ-----------VTPKEVEKWKEDFVRL
Human C3        ---MGPTSGPSLLLLLLTHLPLALGSPMYSIITPNILRLESEETMVLEAHDAQGDVPVTVTVHDFPGKKLVLSSEKTVLTPATNHMGNVTFT------------IPANREFKSEKGRNKF
```

```
Ciona C3-1      TNSNSDFT---ENIDVLVSKQSGYIYVITDRPIYKPNDTVKISAFLLNQNMGHQTG----VDAEITIQTPDGIGLVRESFVDLQS-----NRLNHEFAINENPIYGTWSIEVKFSSDGYT  219
Ciona C3-2      ARATFTFT---KRVQVLIDRNSGYLFVQTDRPIYRPNERVEIRTYPLQQDMSPEKN----AMVQVIVKTPDGIGVNKVERQLPAS-----GFIDTTFHVAEHPMFGTWFVQAKYVSKAFT  136
Halocynthia C3  FGVNCQFQNPAEKFTIPLNSTHGYVFMTTSRPAYRPGENVHISVAGLAQSFTAARGDLPLPDVTIKIRTPISLGRMSVQDTLKRMDNEPTGMLRHTYEIPEDPMTGVWWAEAVMNN--E
Amphioxus C3    DDPQLTFQ---KEAQVLLSYQQGYVFVQTDKPIYTPNQKVKMRIMPLDQDMTPASQP-----VKLEILNPQGIIVERKTFPGSAT-----GFIAETFDFPAFPLFGNWTAIAHYGP-EMQ
Sea urchin C3   LNPIYPFE---EQTDILVTLQSGYVFVQTDKPIYTPNQDVMIKVMSLDQDMLPSDR-----ELHVEIMDPSGTSVKRLRETHADVP----ASIGRLQVHNHPASVNGPFRLSMDP--SS
Hagfish C3      MVKWDGGQ---HMEIDIPLTSRRGLVFAQTDQPIYTPNNDVNIRLFPVTRQLNPILS-----SLVVDIMNPDGVVVDRIEKNAFEVEK---VMELRPFHVPAITSLGDWKIVSWMKD-KPQ
Human C3        VTVQATFGTQVVEKVVLVSLQSGYLFIQTDKTIYTPGSTVLYRIFTVNHKLLPVGR-----TVMVNIENPEGIPVKQDSLSSQNQ----LGVLPLSWDIPELVNMGQWKIRAYYEN-SPQ
                        *       *  *  *  *   *
```

```
Ciona C3-1      TSSTTSFKIDKYVLPTFDVALQLAQ-SHILTSDPRITGTIYANYSYGEPVNGNVYLSATLQKLPGGPAIKFYQIPPRITRTALFRNGVKPFSIPLVELIGILQPGETLAQMSDQGAVVSI  338
Ciona C3-2      TTAEATFAVRKYVVPTFNVQLELER-NYILISDSHITGKIMANYSYGLPVSGNYFLSMKLKRSQNGEPQEFYKMPGNGTLKANFRNGIQRFNVSVNLLLNILNPGETIADLAEMGATFCV  255
Halocynthia C3  VMATTSFVIDKYVLPTFDVKIEMEQ-HFILPGQEIINGEIRAHYSYGEPVDGRYYLSCILKYGIVGTEKEFFKIPKSPFGEAIISGGKKKFSISTSKILEATN-FQSLEDFIALDGKVII
Amphioxus C3    LNVSTQFEVKEYVLPTYGVRIIPSN-PYILPQDDVISGEVEALYTYGKGVDG--FLDVKFGIIDLEGNRQLFAQLQT-----EVNGGFGFYEIDTQRIKDLDL-------WFPEGSRLYL
Sea urchin C3   LSSQVTFVVKEYVLPTFSVSIETP--KYILAGADSVVSTIIAEYVFGKVFG---RYTFKYGVVDNGTVQELGTKQG-------LLGDSGRAEVTLDDLSTDFG---EDWFVRFRGRHFHA
Hagfish C3      FNYTSGFKVEEYVLPTFDVSITSEQ-PYLHVYDKAFTIHIKAMHIYGKPVMGRAYVRYGVKHQSKRTLLSTSSALAR------FEQGEAMHTLRQKHILEQYP-----DPKLLLGQSLYV
Human C3        QVFSTEFEVKEYVLPSFEVIVEPTEKFYYIYNEKGLEVTITARFLYGKKVEG--TAFVIFGIQDGEQRISLPESLKR----IPIEDGSGEVVLSRKVLLDGVQN---LRAEDLVGKSLYV
                ****  *                                               *       *   *  *
```

```
Ciona C3-1      FATVNGEADGVMESAVISNIPILRTPYRIDKSRALKYHTPGISYLLQVDVQDVVTHQNMANIPVRIEITGPN-------------GQRVIRNSTANQNGQVNYPHNFDTTGTQ-FIKVTT  444
Ciona C3-2      EATVNSRADSIMESDVTADLQFLKSPFIIDTSITSKYYIPPVAYTLQGIVTDAISLTPKQDVRIRISVASS------------------TYTTTTNSNGKFVTAFSLVGNGNQ-VVRIQT  356
Halocynthia C3  LATVTGRAQAVYESDMVDDIIFSRTPYIIDMSRSAKFFFPRKIYEIKAIIKDVVSGHPISGVPVTIKAGDA----------------EAISKRSGDNGEIYHAVNLNSQRQQEHILIST
Amphioxus C3    EAAVTEEAGGLREMAVLTSVRFETSPFRIGYDLTATHFKPGLPFLVKLTLTYPDT-KPAQDIPVRVSATAIIPGQDPIVILG---RNNEHNSDTTNQYGQASFTVDVPPGTQTLTVTAKT
Sea urchin C3   NAIVHESATQFSEASINTKAIFVDSPYKFSTKRTVQHFKAGLNLQVKLDLTFANG-DVAPNVPVNVIATATLRDGSEELLRRPDIAGGQEEISNTDDQGGVSMVLNVQSSITSIHIVAAT
Hagfish C3      EASVISSDAGEIENSILDDIPIVASPYSIKSKWTVPFFKPGVPYIYKVLVLNPDG-SPASGVPIKVSFSFDSS----------GNWITQKRKTMDNGIAMQTINTARNSKKLNIKVQT
Human C3        SATVILHSGSDMVQAERSGIPIVTSPYQIHFTKTPKYFKPGMPFDLMVFVTNPDG-SPAYRVPVAVQGEDT-----------------VQSLTQGDGVAKLSINTHPSQKPLSITVRT
                  * *                                                                                    *                      *
```

```
Ciona C3-1      AKPNLGAANQATVNITVEAYNSPGQ--RFLTVTPNPH-----TVDVGREVHITLAFNQPTPAEIRFYVVSRGSVVLVGRIVPPANSH-NIVQIIQVTQAMVPSMRVIAYFL-HGA----E  551
Ciona C3-2      LDPDISNEQQARVNLTIQPYRSPTS--SYLQILASRH-----TVTVRRTFQLTFTFGSSRPTDIRYYVVARGGIVLSSVVRLTPLNR-QKTINVWPSQAMVPFARVVAYYF-KDN----E  463
Halocynthia C3  NHDTYKLEEQSNITLSVDRYDAKT----ELGIQLKKQ-----FVDVGDTVSVELRLGDVSPTDIRYYVVSRGEIVYAQTKEITENG--DVAIDFEVTHKMVPFSRVVAYYF-LNG----E
Amphioxus C3    EQVGLPVAHQAQENFEATPYQSPSGS-YLLVRVLQRG------YPVGPVTVTKQNDIQSYNYMVVTRGPVTLQGKIVRQGGV-LKTITFRTSAVMAPISRLIVYYINLQG----G
Sea urchin C3   HDPQYP-NNQADISFEVSPGVSANGH-DYLVIRPQNTDDRNLNVDVSTDFLIQRIGTTGDQDIDLHFLCITGGKVVLEGVQRSISAA-GTNLAITIRAYMAPQMRLIVYYVTMDG----S
Hagfish C3      EDERLEQSQQAEASFTIASYSSPSG--SFIHLNAHREVK---SPGEHIVFDVFIKSAAKDHVLHFNYLMISNGKIHNFLQEGRKGD--TTSVSLLLTPELVPQFRLVAFFILPSG----E
Human C3        KKQELSEAEQATRTMQALPYSTVGNSNNYLHLSVLRTELRP-GETLNVNFLLRMDRAHEAKIRYYTYLIMNKGRLLKAGRQVREPGQDLVVLPLSITTDFIPSFRLVAYYTLIGASGQRE
                *                                                                                      *                       *  *
```

↓forms disulfide bond with α chain

```
Ciona C3-1      VVSNSAFVNVVNRCETELTVTT---------NRNTVKPGAPITYTIEGAQNADVLLYGVDRAAYFLYNGSRLTRNSMFSDMAAYDQGCVSNGGSDGPNVFFGAGLTLTTPTHKPGALDTL  662
Ciona C3-2      VVSGSLWFDVVDQCKRELSIEV---------AP-LVTPGATFPITIS-APHALVNQDAAYYLYNGSRLTRDVMFKRMESHDQGCVRNGGEDWNHVFMGAGLSLYTSEQNPTIVDSL  572
Halocynthia C3  VTANSAWFDVTDQCLEEITIEP---------QSQTAKPGEDFKFTVSGPANAKIEFSAVDRAAYFVHNSSRLTRNTLWRKMEAYDTGCSRKGGKNAMSVFKEAGLLYQSANINNQDLHTD
Amphioxus C3    VVADSTLLEIENVCRNKVMVS----------STEDVEPQEQANIEVNADPNSLVGLLAVDQAVYLLNMYNRLTSQKMFQAMAKYDQGCGPGGGQDSANVFKDAGVTVLTNTQLSPAVRSN
Sea urchin C3   VIADSLLLGVEEKCRQDPQLSLDILPRNVGPERDVYEPNGQIQVEVTAPIDSNVGLLAVDKAVYLLRDKDRMGKQKMYERMRSYDTGCGPGGGQNTAQIFKDCGMTVLTNAGLDVPIRED
Hagfish C3      LVADSIIIDVKDSCHAKLSLDVAG-------GKRLFSPRDNVNFDLSGESDSWVAVGVVDKAAYVLDKKNKLTANKVYKAMEASDLGCSVGSGKTGPLVFRDAGLAIMAKEISGMDDVKD
Human C3        VVADSVWVDVKDSCVGSLVVKSG----Q--SEDRQPVPGQQMTLKIEGDHGARVVLVAVDKGVFVLNKKNKLTQSKIWDVVEKADIGCTPGSGKDYAGVFSDAGLTFTSSSGQQTAQRAE
                *    *                                                                       **                   * **   *           *
```

β chain ←   → α chain   **CiC3-1 convertase cleavage site↓**

```
Ciona C3-1      DCTAAAQSRKKRQIQVEQEL-----SIEAKLHKCGEDGKKKSFSECDT---CEMQRDRVIYTF---DDAIPGCSERFYAECIALARLNSG-----------TRRQRVQGRS---IGVN--  755
Ciona C3-2      NCDR--NSRNKRNVELEDSVLLT--SLNRKLQQCRRDGKRDAFVN-ET---CEMRTARCGINY---GDQYPGCCEVFHQSCVQASLQNSGNEGEAEASVAQKRASFQGDSSPFAVEVEQP  681
Halocynthia C3  VCAPEGGARKKRAANPNRDFQL---VAAAKVEKCRLDGQRLNLIH-LT---CEQRLAIISANM---QHKDENLRNSCMTAFEEACVALEN---------DFYTQESVSRSLSAN-----
Amphioxus C3    -AGCGGQSRRKRDLQNSLEAKVM--EFNETLQPCCMDGQQWDPLG-RS---CLQRAKLNSTSQDECYFAFLTCCNHARSLRRLGRGR-----------------GRMGGGGGLLDIDID--
Sea urchin C3   -VECMDEDTRRRKSIDRDQLCLY----DPTYLADCLADKPRLRRVLTEGGWLCPRRARVLEAECNLEPDQTQVYRDCCINSMNPV------------------VTATSRSGGDGG---
Hagfish C3      PGCPNGHTRRRKELVLEIAIEKAS-TYPAELRKCCRDAAIESPLR-LS---CEERTKHIHDEGEGCQETFLECCKHVEEELLIAMEEE----------------DEDLGRSQGEDFMIQ-
Human C3        LQCPQPAARRRRSVQLTEKRMDKVGKYPKELRKCCEDGMRENPMR-FS---CQRRTRFISLG-EACKKVFLDCCNYITELRRQHARAS--------------HLGLARSNLDEDII--
                        *                    *  * *  * *              *   *                 *                             * *
```

CiC3-2 convertase cleavage site↓

```
Ciona C3-1      -----GQVERSVASAV-------------------------------------------------------------RKDFAEALYLDP  778
Ciona C3-2      EGAILSQVSRPAVQSLGQSLAPTRLPVFAVQAARPPFLAMAHTFLGMAPQPMAHRFVPNPFILSSEPTPPATTTTTTTTTTTTTTTTMSTTTATLPPVANNPGRERSNFQERLSWPT  801
Halocynthia C3  --------DNANEFRLRK----------------------------------------------------------------RSDFKESWISEE
Amphioxus C3    --------EDESQLVA-----------------------------------------------------------------RTEFPETWIFED
Sea urchin C3   --------EQNAAVKV-----------------------------------------------------------------RDDFRETWFFDV
Hagfish C3      --------ESQVVIR-----------------------------------------------------------------SHFPESFMWEI
Human C3        --------AEENIVS-----------------------------------------------------------------RSEFPESWLWNV
                                                                                                 * *
```

↓ forms disulfide bond with β chain

```
Ciona C3-1      STLEQAGR-----RQITAKARDSITTYEIDAMASADTPDGFCIAPTTNVKVFKNVFVQVYTPYSLKKREQALIKLSVFNY-GDTLVTVDIMMRAHPVLCTHFRTDGSYDLVRTISVGPNS  892
Ciona C3-2      IRIRPNGH------ITSYKTARDSITTFVVGAVGMQDSPDGFCIAPTKEMKVFKDVYCQVAPTKEMKVHKNVFIQLYLPVSVRVREQAVIRFAVFNY-GPQEIEVTVKVLHSKEICTNFIS-DDFRKLVVLTIPSRR  915
Halocynthia C3  KKIGDDGT-----LVFRENAPDSITIYEIAAFGMS-TVSSFCIAKPKEVKVHKNVFIQLYLPVSVRVREQAVIRFAVFNY-GPQEIEVTVKVLHSKEICTNFIS-DDFRKLVVLTIPSRR
Amphioxus C3    VQVDDRGQ-----AVVPVTVPGSITTWVIQAVGIS-TANGMCVAKFPRMKSFKKFFIHLQLPYSIIRGEQVARATIFNY-DQQDLRVNVYMQGVEGVCSGARA-GERSERKTLFIKGND
Sea urchin C3   VSMPEDGSP----YLYPVSIPSSITDWHLTAVSLS-PTQGMCVEDETTVSFVQDFFIQLHLPYSVVRLEQTQVIATIFNY-GFSDFEVSVNFTVDQGLCTAENP----AIRHVFVESKR
Hagfish C3      IKLSRSAENG---KSRITKKMPDSITTWDIQAVEVS-QSKGLCVGPSLELTVFKQFFLKVHTPYALKQYEQVELRVVIYNY-MNQDVKGEIQVKCGDGICTDAEQ---NEPLKSRFAVEKNS
Human C3        EDLKEPPKNGISTKLMNIFLKDSITTWEILAVSMS-DKKGICVADPFEVTVMQDFFIDLRLPYSVVRNEQVEIRAVLYNYRQNQELKVRVELLHNPAFCSLATT---KRRHQQTVTIPPKS
                                   ***    ***  **            **          *   *    *  **       **     **    **
```

```
Ciona C3-1      AGSASFAVLPLRIPAGDIPKATVEVYI---------TDNRNRTYDSVKKEIL[IEDEGELKDIYET----------FPIDLKNSR]VQQTQINFTFPEQFVLG-TRKCMLYAYMDFMGPAIE  992
Ciona C3-2      FASAPLTVIPLLIPLTGRSPIQLKV-V---------NDQTNVIQDSIRRVVLEPAGEMKDTYNS----------YPIDLSTG-NQSIEIQLNFPEQINLE-SRKCWIYAYASYMGPSIE  1013
Halocynthia C3  SATSYFTILPVKIPKGGKSKIELKA-----------EGDSSENSDWIEKMLLVEPPGQHNDSHGT----------FLLSPDVN-GGVIDFEIDLATKNVLAGTIKCNLYFYGHIMGPNIE
Amphioxus C3    AASVLFPIIPLEVGTFPIRVVAFSTAAGGDIIEKSLQVIPNGVEERRLVRSIFVDPKGRARDRKREGEEEVALPTEHDVVPDDNG-LQFDVVDVRLPPETIEG-SEQCAVSIMGDIMGPTIT
Sea urchin C3   AASTTFHCTPCGSRGIPYTVTAGGSKKRDSVRN-NLRVVSQGVMQRKSRSLTLNPGRVMFSDDVTTPSPNNSLGSGPGTIFIGEGFQHEEIAISLPGSSIPD-TESCSVKLIGNLLGTAST
Hagfish C3      ATSFSFMVVPLSSSDSSVSVLARVFG--------------SDVHDAVEKDLRVMPGENYEEMSRS---------WSVQPRRHGGQQVIVVDNETPQNVVPGTEMSAFLSAQGNLVAETIQ
Human C3        SLSVPYVIVPLKTGLQEVEVKAAVY--------------HHFISDGVRKSLKVVPEGIRMNKTVAVRT------LDPERLGREGVQKEDIPPADLSDQVPDTESETRILLQGTPVAQMTE
                *
```

thioester site

```
Ciona C3-1      VDPVTQEANNVNSLIRQPYGCGEQNTMIYAGRTMIYAGRFVFALQYLVTTGTITPNS-P-EYNSAVNKIEAAFQREMTYRTTHTNPRVMSVFTHYLPSTWLNAFVDKVFYHGRRYDTDMNVGPICNSL  1110
Ciona C3-2      VNQITQEPRSIASIFRQPYGCGEQNMLVTGPNVYAHMFLVTTGKMAPGS-I-RYEQSTRRMEDGFNQQMRYRSDLVGKRAWSVFSHYRPSTWLNAYVDRVFIQAQVYYTEMDLTPVCRSL  1131
Halocynthia C3  VELEGRLAN-LQNLINSPGGCGEQNMIRIAPVVYIHAYRSNLEAFTVTD-A-QRAQTLKYIQDGYAHELEYKTQVPQGWAFAVWANNPPCSTWLNGFVSRVFASARKYWPGMEVDRICQSV
Amphioxus C3    TTIGG-----LGTLLRLPTGCGEQNMIKLAPNVYVLSYLHCTDQITKD----VEERAYDIYADGVARELTYRQDN----GAYAAWKHRPGSTWLTAFVVKVFCQANRFTR-VEPGHVEGSI
Sea urchin C3   DPIGG-----LDHLVRQPRGCGEQNMYLAPTLFVYQYLIAVGSDTAE----QEARIYDYIADGVARELTYRQDN----GAYAAWKHRPGSTWLTAFVVKVSFQANRFTR-VEPGHVEGSI
Hagfish C3      NTLKGSK---ISNLLRLRLPTGCGEQNMMYSITVMVARYLNRSDQWNKMGDPQLKKRSFDFITSGFASQLTYRKPD---YSYAAWLHRASSTWLTAFVVKAVFSQARQLVF-IPVSEICGSV
Human C3        DAVDAER---LKHLIVTPSGCGEQNMIGMTPTVIAVHYLDETEQWEKFG-LEKRQGALELIKKGYTQQLAFRQPS---SAFAAFVKRAPSTWLTAYVVKVFSLAVNLIA-IDSQVLCGAV
                                *  *****  *                                                                     ****  *  **
```

Fig. 1 Legend see page 1059

```
                         catalytic histidine ↓
Ciona C3-1     NFLIGE-QMAGEHFRERRPPLHREMHGAVKG----PMTLTAHVAISMGEINSICLPELNQRVIASRVSAMNYLEQHKDHATFQRPYPLSLLAYAAALHNPRSQLAIEMNARLMAMKQTSG 1225
Ciona C3-2     QWLVGE-QHQEGYFLERSPVIHREMHGAVGG----RYSLTAYVLVTLIEAQRINCSGVNQQIQQSRDKAINYLRNNRNHPAFQRPYGLSILTYAMALHDQSSAFTIELNQRLLGFQQLDD 1246
Halocynthia C3 AWLLTQ-QEPDGHFDEDDPVHHKEMDGQVTN----DITMTAFICISLYES-RSSCPELSTRIQEGVNSAMEYLKNHFDDAKNT-PYTVAIVAYAFAIWEPHGQFAQKWNEQLVAMKIEDE
Amphioxus C3   EWLLST-QREDGAFKEVYKVHHREMTGGVQG----DASMTAFVLISLLEN--CECPIAERSIAIER----ATLFLERQLEQLKRPYVIAIVTYALHLADS--PLKGAANEKLRSIAKYDE
Sea urchin C3  NWLIDNNQLPSGAFQESQQVIHQEMIGAVKG----ETSMTAFVLISLLES-RNLLVPANQKIDEAIGKATEYLVT--QVENIDRVYDKALVTSALRESAS----VGTANGKLWEDRNEDG
Hagfish C3     RWLMRK-QDKDGSFLESKPVVHLNMMGQVTG----KVVLTSFVFIALLEA-RESCINEVEGFTVVVEKAHGYLTS--QAMNGLEDFPLAITAYALSLWKVSDGAAKVTMHTLKTSGLQTE
Human C3       KWLILEKQKPDGVFQEDAPVIHQEMIGGLRNNNEKDMALTAFVLISLQEA-KDICEEQVNSLPGSITKAGDFLEAN-YMNLQRSYTVAIAGYALAQMGRL--KGPLLNKFLTTAKDKN-
                 *       * *    * *  **       *                 *                  *      *  *                      *             *
```

```
Ciona C3-1     NGAYVFWRAKTLAEISGTNAHAYWYRTRPLALDIETTSYALLAQVAMSKAKDPLVQADLDLGRKISLWLISQRNSNGGFRSTQDTVIGLQALSTILSWISTVEPGIANPRIDFEMEGAAG 1345
Ciona C3-2     N-SYVHWQAHSHSDIQGTDTHDYWYVRRPQAIDVETSAYALLAQVKLASSNNNVQQSDLDTTRSIALWLISQRNEAGAFRSTQDTVVGLQALAEYMAWVSQVQPQSADVNIRIQLRGPPN 1365
Halocynthia C3 K------------------FVHWQGTRGRAADIETTAYGLLAEISLKK----VTENNYRWTKQIANWLISQRNDRGGFISTQDTVVAIEALTKYMTMLETAFQK-ENVDLRVILSSTSQ
Amphioxus C3   GTNSRYWEADASSIADGQ--QPYWYTRKPSAIAVETTAYALLTQMHIG---------DIQYSNPIVVWLTQQRNSAGGFVSTQDTVVALQALSSYCG-STKVDPTQFTCQMTSDNDLDYN
Sea urchin C3  TG-AVSFTPDDANYEDGS--QPFWLQRKPSAIEIETSGYALLAQLALL---------DYQKAGKIALWLSKQQNDGGGFVSPQDTVVALQALAKYTE-RPEFNTIEMNCDVATEQIPLHR
Hagfish C3     E------------------LIHWGSNKGKAAAVESTAYGLLAAIQHE---------EGEIAEKATNWLSQSATFGGYFQSTQDTVMALQALTGFESCQSRMKKMDLSFKIRAEENGVFD
Human C3       R--------------------WEDPGKQLYNVEATSYALLALLQLK--------DFDFVPPVVRWLNEQRYYGGGYGSTQATFMVFQALAQYQKDAPDHQELNLDVSLQLPSRS-SK
                      *           *     * **              **        **      *   * * * **
```

```
Ciona C3-1     SAWAPPNAYTNHIDNTNQRLKKQMAVPKNVIEGTIVDVTSRGQGQGILSYRC-TYRTVVDEESCHYNLSTNVRVMNPG----------------------------------------- 1422
Ciona C3-2     VAWSPSGDHIMNLDGQNVGFRTEVEVPAHIIGRGNIRAHISGSGEGIISNRC-IYRTSSAEGACHFNINHTVDFVPSPDGRV------------------------------------- 1446
Halocynthia C3 HSSWFENPEVVGITENDGMTRKTVEVPKDAIDKP-IRARVDGTGEGSLSYRC-TWRQYESADKCHFQVEVSVDEAQEE-------------------------------------------
Amphioxus C3   EEIHVDKDNALINQEKTAPVGGKLFLSTSGTGIGQMQVEVRYHTPDVHRERC-LFEVVVTTEEAEGPVEPDEEPEEGQEGDEDYPDYDGELDARERSAFSSRSRIGGLLSSGRSRSRIAR
Sea urchin C3  YHIGNDNAKVQEEVDVSPSIGRSLTFDSRGTGVAKANVELRYNTEKSDIDTC-PFHLNISAVEVPSDEVSGQNVKC-----------------------------------------
Hagfish C3     KEFQITNDNAFVQKPFKVPVHGQLTVTASGTGQGILTFVKKYREKVVIKKDCKGFSLEITTNLDNQVKQRRRQSINPEFN-----------------------------------
Human C3       ITHRIHWESASLLRSEETKENEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCNKFDLKVTIKPAPETEKRPQDAKN--------------------------------------
                                                               *
```

```
Ciona C3-1     ----VQGKRIKITITVSKPLGPQ--SEATILYVGFMSGYKAVDQTLEDLSIDENVAGVVDRYEVT-----NHNVLFYLNRITS-QSMNISFEMRQMVLVRKPQPGKINIYEYYEPVVHCG 1530
Ciona C3-2     ----AEELPVKLRLTISKQVGSP--AEASIVEVNLMSGFSVVENDLQRIAISEAVDGLVDRYELT-----SSKVVLYLRRIGA-EPTTLSFRMRQTVVVSKPQPAKINVYDYYEPSISCG 1554
Halocynthia C3 --------DMKVVKMTVSKNGTK--AQMSIIDVVMLSGFEADKDSIAALNKDIAADGIFDRYELT-----NNAVKFYLGSIGE-KEIKFAFKIKQVSVVSKIQPASVIVYDYYEPDIRCT
Amphioxus C3   QAEDDQSQHFYLVRVCTSYRGQRGASNMAIMDIGMFSGFEPVKQDLEQLLN-RAGKKAVQRYETT-----NRAVLLYFDEITS-EEICVSIRMKRVMLVGAVQPVPVSVYDYYQPDEACT
Sea urchin C3  ---------LMITVCTSYNGDG-TTHMIIVDVGLYSGFKAVEEGLTGLKQVVNGSSLISSYEAS-----SRSVIFYLDTIPSDEDLCFITFSAESDVVVGNVQAAAVHVYDYYDPEKSCT
Hagfish C3     ---------VYRFIGCFRYLRNQ-EPGMVVMDISLPTGFEAKKKDLDDMKN--LVDNYIVQYEIR-----PGRVFLYLDKVNKDEKNCVGFRLNQVFESNLVLPVTATVFEYYEPDFRCS
Human C3       ---------TMILEICTRYRGDQ-DATMSILDISMMTGFAPDTDDLKQLAN--GVDRYISKYELDKAFSDRNTLIIYLDKVSHSEDDCLAFKVHQYFNVELIQPGAVKVYAYYNLEESCT
                                              **              **          **              *        *                            **    *
```

```
Ciona C3-1     QFYSLPDTTPDLQTSNCETGANNDDNLCRCAEGGCPTCRTRDDTLFKLHVLEV--EVGARCGTCTQNR-CYNHRDEGCSAKYVYKVRVIGVREDS---VFVNLDAVIEDVYRSGSD--NV 1641
Ciona C3-2     IFYSLPEGVDRLRTTQACTSGNTN-GLCKCAEGSCPVCRTREEQLHNTTCLG----SSGQDCLICQGGS-CENMYRKACTSNYVYIATVTSQPRDAWTPGYTSFSVRVDEVIKEGNDEISV 1669
Halocynthia C3 KFYTVGDPNLKLRTICEDGG-----VLCKCAEGDCPVCRSQNDQLTNTQCSSHGNHEQICSDMCEDNIGCYNLHVETCKRDYVYIIKVESVEETE-GSGYKIFKARIVEVIRSTSKHIDP
Amphioxus C3   TFYHPGQGSPLLATLCDGSQ-------CVCAEGKCPKETSPK-KLRELTQRD------LEEKACN-------------DLDYAFRVIVMEVRQEG---SFDRISMIVESPIKKGLD--DV
Sea urchin C3  IFYKPGDGSALLSTLCSENE-------CICSGGSLEYCNIDPCPGPYTRID-------LEGTACAS------------HSSYALKIRIDEVEIKEG---FRICKFTVLNPIKTGDE--DV
Hagfish C3     KSYHPKMEVNPDASCHGN--------ICNCLQRHCVELKGMADEDRNADRN---------GNACR-------------AEYVFIIGVTKVTKTAS---YININAALKTVLKKGMD-QAI
Human C3       RFYHPEKEDGKLNKLCRDEL-------CRCAEENCFIQKSDDKVTLEERLD----------KACEP-----------GVDYVYKTRLVKVQLSN---DFDEYIMAIEQTIKSGSD--EV
                *             *            *         **                 *                   *           *                    *
```

```
Ciona C3-1     QNGTVRTFVIRNDCHASCRDTDEDLRLLRNNIPAKDTRLFKEGRVLLITGSLIERRQNRNRLLQTVYQVDEQTTAERLVTDVACARAKSVLVR----CAG----NRPPRQSKCVKSRELN 1754
Ciona C3-2     DRNRIRQFTTRKECYEKCKDP----VLAGPDRSEESQRYLKPNMKILLMSNFLDFSMDRSGHVRHDYQMGEGTTVERIIPDSKCVQIRARVALPKFQCENPNFHRNPNRQAKCDKMMKMK 1785
Halocynthia C3 KDD-VTSFRIREDCYAESCKDPRFEEIARKKNNAVEAKFIEEGRHLLVYGSNPTVTEAKNEPVQYIYELDDNARLEIVPSGEKCSPEILECAP-----------NDIRCNKKKAKLDKIT
Amphioxus C3   FDGEERVFWKRKTCSG-------------------LQLVEG-TTYLLMGKDGTKYTDEQGFDSFRYVITEQSFVAEWPTG----------------------------AKAERKKFK
Sea urchin C3  PHQAQRQLFINEGCDCP-----------------KVKGRNIGGTFLLVGQKSLKYTTEQGEERYRYVYGPTSKLEFWPATR----------------------------KARNAAVFD
Hagfish C3     NVGARRSFVIPMHCGKN------------------LNVSPGDIYLVMGMHNAHWRNSD---RTQYVLTSDTWFEKFPLES----------------------------VCRLPSPP
Human C3       QVGQQRTFISPIKCREA------------------LKLEEKKHYLMWGLSSDFWGEKP---NLSYIIGKDTWVEHWPEE----------------------------DECQDEENQ
                    *                               *                   *                  *                              *
```

```
Ciona C3-1     AICENMKRLKESLNDIGCD- 1773
Ciona C3-2     TSCDNMDRLKN-QVQQGCDK 1804
Halocynthia C3 KLCKINSTLIE-AHLEGCDV
Amphioxus C3   RVADKFQQLADKIFLQGCTT
Sea urchin C3  KLVAFEAQMAPASACSNEN-
Hagfish C3     ASCQVSENFKG-CSLKG---
Human C3       KQCQDLGAFTESMVVFGCPN
```

Fig. 1 Alignment of CiC3-1 and CiC3-2 proteins with C3 proteins from other species. CiC3-1 and CiC3-2 were aligned with other C3 proteins using Clustal X (ver 1.8), with a few adjustments made by hand. Several important functional domains and amino acids are *indicated*. The *underlined* sequence is the putative leader sequence of CiC3-1. The two sequences *in the box* correspond to the peptides used to prepare antibodies

this sequence insertion in a protein database did not yield any significant similarity to any other protein. No α-γ processing signals were found in either sequence. The predicted relative molecular masses (M_r) before processing were about 198,000 for CiC3-1 and 202,000 for CiC3-2 and the deduced molecular masses of the α and β chains of CiC3-1 were about 124,000 and 71,000 respectively. The deduced molecular mass of the CiC3-2 α chain was about 137,000. Putative N-glycosylation sites were identified in CiC3-1 and CiC3-2, five located in the α and five in the β chain of CiC3-1, six and five located in the α and β chains of CiC3-2, respectively.

A canonical thioester site (GCGEQ) is present at positions 1,012 and 1,033 of CiC3-1 and CiC3-2, respectively, and a catalytic His, involved in thioester-binding to the pathogen surface, is conserved at positions 1,131 and 1,152 of the two C3-like molecules. In vitro mutagenesis studies showed that the presence of His[1,126] (human C3 numbering) determines the specificity of the thioester for hydroxyl rather than amino groups, as a result of the formation of an intramolecular acyl-imidazole bond (Law and Dodds 1997). This residue is conserved in all C3 and C4 molecules, except for cobra venom factor, human C4A and trout C3-4; and, in these proteins, His is substituted by Ser, Asp or Thr, respectively.

In addition to His, the Glu residue located two amino acids downstream, corresponding to Glu[1,128] in human C3, is also conserved in CiC3-1 and CiC3-2. This residue forms a hydrogen bond with His[1,131] and His[1,152] and, in humans, is thought to render His a stronger nucleophile, thus increasing the rate of acyl-imidazole formation in C3 and promoting its specificity for hydroxyl nucleophiles, as compared with human C4B (H[1,126], Ser[1,128]; Nagar et al. 1998). C4B shows a considerable reactivity with both amino- and hydroxyl nucleophiles. In trout C3 isoforms, the corresponding residues contributing to the specificity of the thioester are His and Glu (in C3-1), His and Thr (in C3-3) and Thr and Ser (in C3-4); and it has been suggested that these residues may be responsible for the differences in binding-specificity to various surfaces that have been observed among these trout isoforms (Sunyer et al. 1998; Zarkadis et al. 2001).
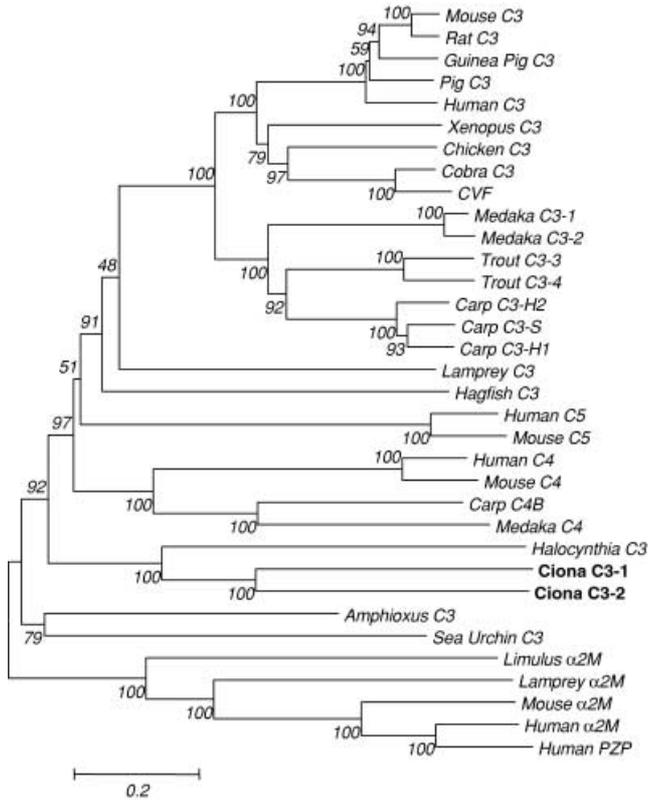
Fig. 2 Phylogenetic tree showing the relationship of *Ciona* C3s to other thioester-containing proteins. The relationships among *Amphioxus* C3, carp C3-H1, carp C3-H2, carp C3-S, carp C4-B, chicken C3, cobra C3, cobra venom factor (CVF), guinea pig C3, hagfish C3, *Halocynthia* C3, *Limulus* α_2 M, human C3, human C4A, human C5, human PZP, human α_2M, lamprey C3, lamprey α_2M, medaka C3-1, medaka C3-2, medaka C4, mouse C3, mouse C4, mouse C5, mouse α_2M, pig C3, rat C3, sea urchin C3, trout C3-3, trout C3-4, *Xenopus* C3 and *Ciona* C3-1 and C3-2 were analyzed by the neighbor-joining method, based on the alignment of the deduced amino acid sequences performed with CLUSTAL W software. The *numbers on the branches* show the percentage recovery in 1,000 bootstrap replications

Fig. 3a, b Southern blot analysis of *CiC3-1* and *CiC3-2*. Sperm DNAs (8 µg), isolated from three different animals, were digested with *Eco*RI, *Hin*dIII, *Kpn*I and *Pst*I, run on an agarose gel and transferred to Nylon membrane. Hybridization was carried out with probes corresponding to the nucleotide fragment 3,955–4,781 for *CiC3-1* (**a**) and the nucleotide fragment 3,872–4,576 for *CiC3-2* (**b**)

Also, the two Pro residues, which contribute to the hydrophobicity in the thioester region (Isaac and Isenman 1992), have a conserved localization in CiC3-1, while CiC3-2 has only one Pro residue conserved.

Our analysis of the deduced amino acid sequences revealed that the total number of Cys residues is 32 for CiC3-1 and 35 for CiC3-2, with 4 and 5, respectively, located in the β-chain and with 28 and 30 in the α-chain. The alignment with other C3 molecules produced evidence that 3 Cys residues in the β-chain and 16 in the α-chain of the two CiC3-like molecules aligned with human and other vertebrate C3 molecules, in which the position of 27 Cys residues is conserved. The Cys involved in the α-β chain junction in CiC3-1 and CiC3-2 aligned with all of the other C3 molecules sequenced thus far. Three Cys residues in CiC3-1 and four in CiC3-2, out of the five present in the C3a region, had a conserved position when compared with human C3.

Also, the C3 convertase cleavage site (Arg, Ser), which is generally conserved in most of the C3 proteins,
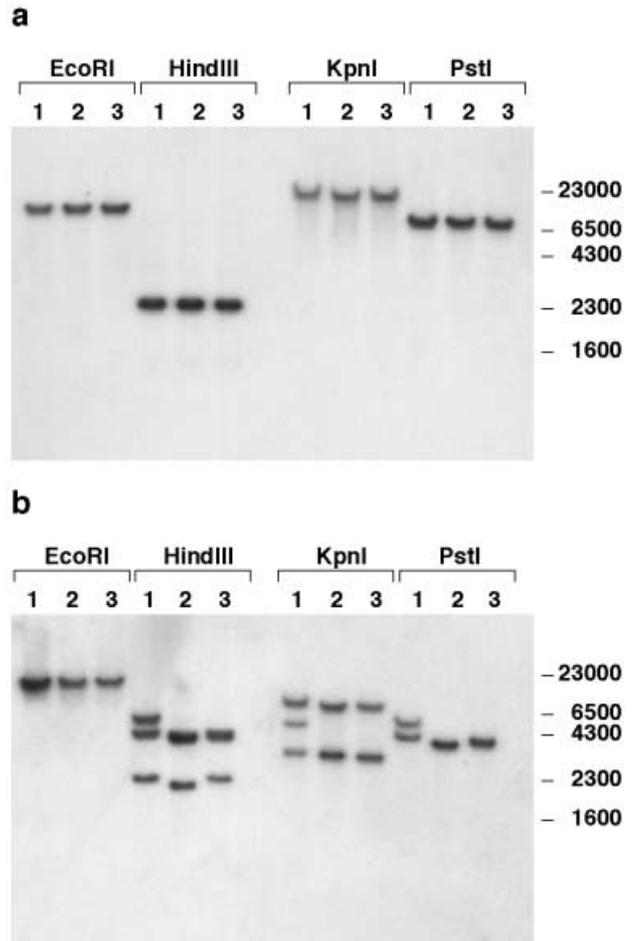
is present in conserved position in CiC3-1 and immediately after the insertion in CiC3-2 (Fig. 1).

To construct a phylogenetic tree, we first used the Clustal software to align the deduced amino acid sequences of CiC3-1 and CiC3-2 with those of C3 molecules from other species. A phylogenetic tree generated from all the available C3 sequences indicated that *Ciona* C3s formed a cluster with the ascidian *H. roretzi* C3 and that the duplication event from which *CiC3-1* and *CiC3-2* genes originate, occurred in the urochordate lineage after the separation of the *Halocynthia* and *Ciona* ancestor (Fig. 2).

Southern blotting analysis

To determine whether *CiC3-1* and *CiC3-2* were present as single- or multiple-copy genes per haploid genome, we performed a Southern blot analysis of genomic DNA from three animals' sperm. Each sample was digested
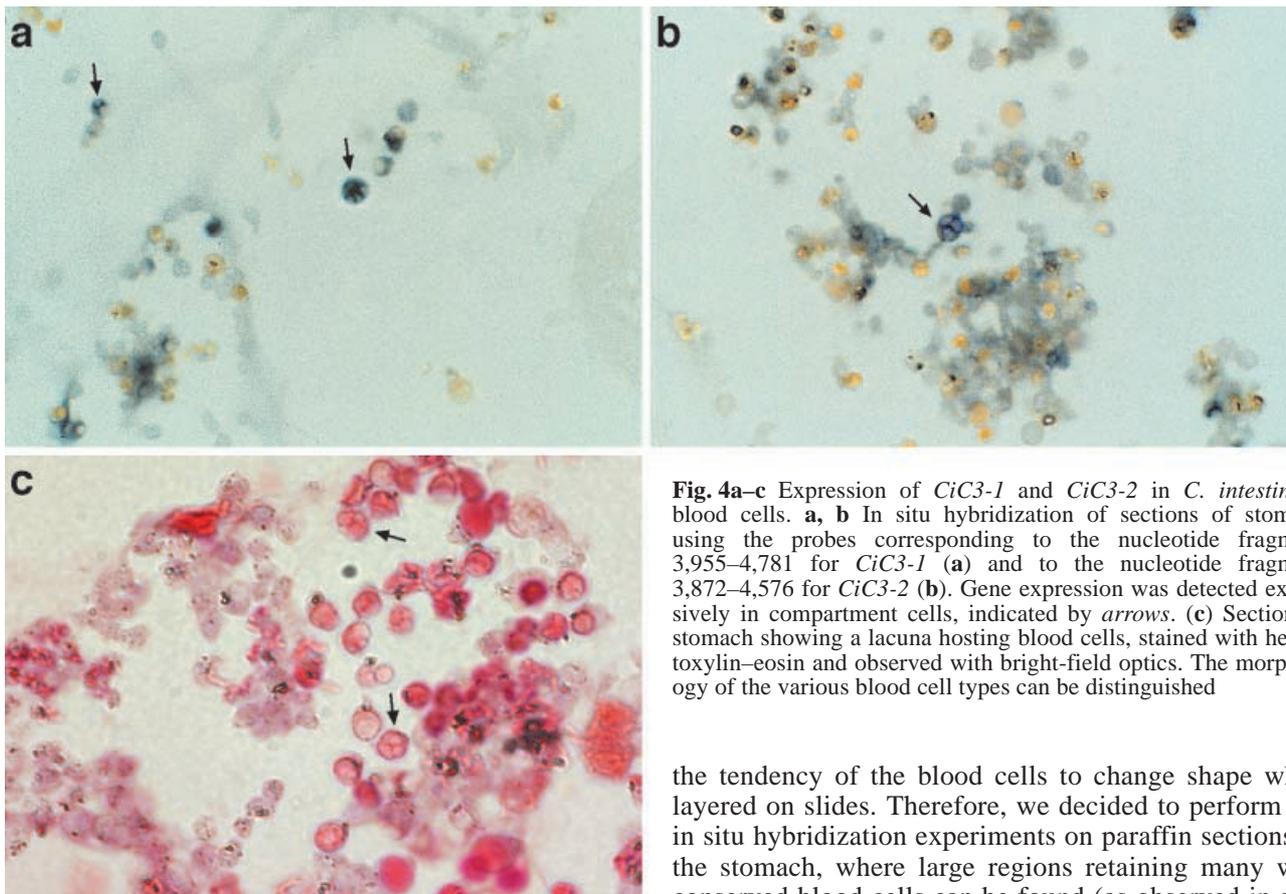
Fig. 4a–c Expression of *CiC3-1* and *CiC3-2* in *C. intestinalis* blood cells. **a, b** In situ hybridization of sections of stomach using the probes corresponding to the nucleotide fragment 3,955–4,781 for *CiC3-1* (**a**) and to the nucleotide fragment 3,872–4,576 for *CiC3-2* (**b**). Gene expression was detected exclusively in compartment cells, indicated by *arrows*. (**c**) Section of stomach showing a lacuna hosting blood cells, stained with hematoxylin–eosin and observed with bright-field optics. The morphology of the various blood cell types can be distinguished

with four endonucleases (*Eco*RI, *Hin*dIII, *Kpn*I, *Pst*I) and divided into two aliquots for use in two identical genome blots. The blots were hybridized with probes corresponding to an α-chain region of *CiC3-1* and *CiC3-2* (an 826-bp fragment for *CiC3-1*, a 704-bp fragment for *CiC3-2*). Only a single hybridizing band for each lane was detected with the probe generated from *CiC3-1*, thus allowing us to conclude that this gene is present as a single copy per haploid genome (Fig. 3a). Digestion with the same panel of restriction enzymes and hybridization with the *CiC3-2* probe revealed one or two bands per lane and three bands with *Hin*dIII and *Pst*I for a DNA sample from animal number 1, suggesting that *CiC3-2* is present as a single copy gene per haploid genome and that the multiple bands reflect allelic polymorphism (Fig. 3b).

In situ hybridization

In *C. intestinalis* blood, according to the classification of Rowley and co-workers (1984), four main cell types can be recognized: stem cells, amebocytes, vacuolated cells and pigment cells. To determine which blood cell types express *CiC3-1* and *CiC3-2*, preliminary in situ hybridization experiments were carried out on blood cells layered on slides. Both genes seemed to be expressed by only a limited number of cells, whose unequivocal identification was rather difficult because of

the tendency of the blood cells to change shape when layered on slides. Therefore, we decided to perform the in situ hybridization experiments on paraffin sections of the stomach, where large regions retaining many well conserved blood cells can be found (as observed in sections stained with hematoxylin–eosin; Fig. 4c). For in situ hybridization, riboprobes corresponding to the nucleotide fragment 3,955–4,781 for *CiC3-1* (Fig. 4a) and to the nucleotide fragment 3,872–4,576 for *CiC3-2* (Fig. 4b) were used. No labeling was seen in the wall of the stomach and hybridization signals appeared only in the lacunae hosting blood cells, restricted to a few cells that can be identified as compartment cells (Fig. 4a, b). Labeling was confined to the bristles of cytoplasm that surround the large vacuoles. The compartment cells have been assigned to the heterogeneous group of vacuolated cells, including the morula cells, compartment cells and signet ring cells; and they account in *C. intestinalis* for only 5% of the total cell population (for a review see Rowley et al. 1984).

No labeling was detected in control experiments carried out in parallel with the corresponding sense cRNAs.

Western blot analysis

To determine the chain structure of *Ciona* C3s, we generated anti-peptide antibodies against peptides corresponding to the deduced amino acid sequences and then analyzed the reactivity of these antibodies with serum proteins by Western blotting. Both anti-CiC3-1 and anti-CiC3-2 antibodies reacted with a protein with an M_r >220,000 under non-reducing conditions (data not shown) and with a protein of about M_r 150,000 (α chain)
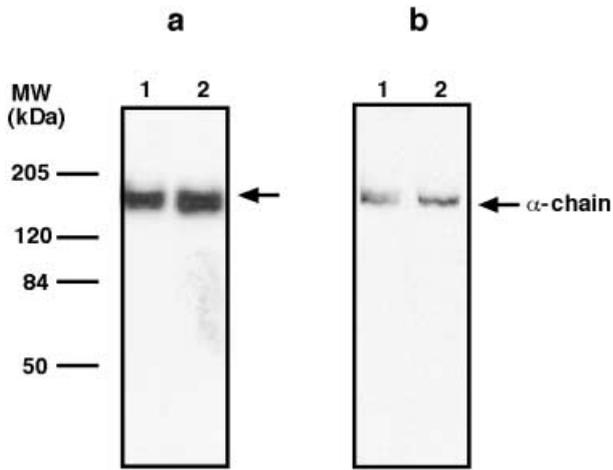
Fig. 5a, b Reactivity of anti-peptide antibodies with *C. intestinalis* blood serum proteins. Sera from two different animals (*lanes 1, 2*) were subjected to electrophoresis on a 6% SDS polyacrylamide gel under reducing conditions, blotted onto polyvinylidene difluoride membranes and then probed with anti-CiC3-1 (**a**) or anti-CiC3-2 (**b**) affinity-purified antibodies, followed by horseradish peroxidase-conjugated anti-rabbit immunoglobulin and enhanced chemiluminescence. On the left, molecular weight (*MW*) markers

under reducing conditions (Fig. 5a: anti-CiC3-1; Fig. 5b: anti-CiC3-2). Analysis of several *Ciona* sera by Western blotting showed the same pattern of reactivity. Figure 5 shows the results of this analysis carried out on blood serum from two different animals (lanes 1, 2).

Reactivity with anti-CiC3-1 was inhibited by the CiC3-1$^{936-957}$ antigenic peptide but not by that for CiC3-2$^{579-599}$. Similarly, the reactivity of the anti-CiC3-2 with the M_r >220,000 *Ciona* serum protein was inhibited by the CiC3-2$^{579-599}$ antigenic peptide but not by that for CiC3-1$^{936-957}$. Despite the small differences in the calculated molecular mass of CiC3-1 and CiC3-2, their mobility in SDS-PAGE was the same, perhaps because of glycosylation differences or differences in other post-translational modifications. Several attempts to purify CiC3-1 and CiC3-2 to homogeneity (to allow us to characterize these proteins biochemically) were unsuccessful because of the low abundance of these proteins in serum. However, Western blot analysis of CiC3-1 and CiC3-2 proteins partially purified by anion exchange chromatography indicated that the generated antibodies do not cross-react.

Discussion

There is a general consensus that deuterostome invertebrates are endowed with only an innate immune response and that adaptive immunity is an exclusive feature of jawed vertebrates. Recent reports have demonstrated the presence of a simple complement system in several species of deuterostome invertebrates. Genes encoding complement components have been isolated and characterized from the sea urchin *S. purpuratus* (Al-Sharif et al.

1998; Smith et al. 1998), the urochordate *H. roretzi* (Ji et al. 1997; Miyazawa et al. 2001; Nonaka et al. 1999) and the cephalochordate *Branchiostoma belcherii*. Thus far, a phagocytic activity related to the complement system has been demonstrated only in *H. roretzi* (Nonaka et al. 1999), while different humoral opsonic activities mediated by lectins have been demonstrated in many urochordates (Ballarin et al. 1999; Coombe et al. 1984; Kelly et al. 1992; Pearce et al. 2001).

Data reported in this paper point to the solitary ascidian *C. intestinalis* as a suitable model for further investigating molecular and functional aspects of the complement system. Because of its phylogenetic position, its worldwide distribution at temperate latitudes, its simple anatomy and the broad scientific background available on this species, it has recently been selected for studies of its complete genomic organization.

From the viewpoint of innate immunity, information on this species is restricted to morphological and functional data. In particular, a primordial blood cell-mediated immune rejection has been observed in *Ciona* in response to allograft transplantation, associated with an inflammatory reaction mediated by lymphocyte-like infiltration (Lakshma Reddy et al. 1975). Only fragmentary data are available concerning the cytotoxicity and phagocytic activity of *Ciona* blood cells (De Leo 1992) and no information is available concerning the molecular basis of any of these immune reactions.

In this paper, we describe the isolation of two genes, *CiC3-1* and *CiC3-2*, which encode two distinct thioester-containing molecules that, on the basis of their sequence similarity, overall protein structure and phylogenetic analysis, are apparently homologous to the complement component C3 identified in many other species. Southern blot analysis indicated that *CiC3-1* and *CiC3-2* are single-copy genes and the restriction pattern obtained for one individual seemed to point to the presence of allelic variants of *CiC3-2*. cDNA sequence analysis of the two genes and the deduced amino acid sequences showed the presence of a β-α processing site and a canonical thioester site in the α chain of both peptides. Also, the catalytic His, the Glu forming a hydrogen bond with this His and the surrounding hydrophobic region were all conserved. The α and β chains are held together by a disulfide bond, as indicated by the conserved positions of the involved Cys residues.

In mammals, the *C3* gene is mainly expressed in the liver by hepatocytes (Alper et al. 1969), but macrophages and other cell types have also been identified as sites of production of C3 (Lambris 1988). In the ascidian *H. roretzi*, the hepatopancreas and blood cells have been identified as the sites of *C3* gene expression (Nonaka et al. 1999). The ascidian *C. intestinalis* does not have a hepatopancreas or any analogous organ in the abdominal region. However, like all urochordates, *Ciona* does have a well-developed hemocelic blood vascular system, consisting of the heart and associated blood channels without an endothelial lining, an anatomic organization which allows the blood cells to penetrate directly into the

surrounding tissues. This organization is particularly evident in the gastrointestinal tract, on which we conducted our in situ hybridization experiments. Analysis using riboprobes corresponding to 3′-terminal sequences of *CiC3-1* and *CiC3-2* revealed the specific expression of these two genes in only one type of blood cell, which we were able to identify as compartment cells. On the basis of morphological data, this cell type is characterized by the presence of few cytoplasmic vacuoles and is apparently related to the highly vacuolated morula cells. The role of both cell types, which probably represent two different functional stages of cells belonging to the same lineage, is at present unclear, although their involvement in the immune response has been hypothesized (De Leo 1992; Rowley et al. 1984).

The presence in the blood serum of the products of both the *CiC3-1* and *CiC3-2* genes was documented by Western blot analysis, using a polyclonal antibody for each molecule, raised against a synthetic peptide derived from the deduced amino acid sequence of a different region in the α chain of each molecule. This analysis, carried out under both reducing and non-reducing conditions, showed positive signals, with a band of M_r >220,000 in non-reduced samples and a band of about M_r 150,000, corresponding to the α chain, in reduced samples. The bands detected by both antibodies showed molecular sizes slightly higher than expected, probably because of the presence of N-linked carbohydrate residues.

In the ascidian *H. roretzi*, a unique C3/C4/C5-like molecule belonging to the α_2-macroglobulin superfamily has been isolated; and its functional involvement in a very simple complement pathway has been demonstrated (Nonaka et al. 1999). It has been suggested that this molecule evolved directly from an ancestral precursor and that the component multiplicity present in the complement system of higher vertebrates could be the result of gene duplication events at an early stage of vertebrate evolution, when wide genome duplication occurred. The data reported in this paper clearly demonstrate the presence of two *C3*-like genes in the urochordate *C. intestinalis* and indicate the presence of their transcription products in blood cells and the encoded proteins in the body fluid. Thus, we can now trace back to deuterostome invertebrates a gene duplication event generating a new complement component that could increase the complexity of the system. Multiple forms of functionally active C3 that are the products of different genes have previously been found only in teleost fish (Sunyer et al. 1998). This characteristic is not exclusive to tetraploid species, however, since C3 genes in some diploid fish have also been found to be closely linked, thus indicating a tandem gene duplication (Kuroda et al. 2000). It has been demonstrated that these C3 isoforms exhibit differing specificities for various complement-activating surfaces, thereby providing a means of expanding the immune defense repertoire.

To our knowledge, the presence of a C4 molecule has not been reported to date in any deuterostome invertebrate. The absence of an α-γ chain-processing site (typi-

cal of C4) and the presence of a conserved thioester site (absent from C5) in the genes we have isolated provide support for our conclusion that the two genes we have isolated are *C3*-like genes. These two genes, which show the same pattern of expression, could exhibit a different specificity toward different activating surfaces, as found for fish C3 (Sunyer et al. 1998). Structural and functional analysis of the two *Ciona* C3-like proteins is in progress to characterize these two C3s.

References

Alper CA, Johnson AM, Birtch AG, Moore FD (1969) Human C3: evidence for the liver as the primary site of synthesis. Science 163:286–288

Al-Sharif WZ, Sunyer JO, Lambris JD, Smith LC (1998) Sea urchin coelomocytes specifically express a homologue of the complement component C3. J Immunol 160:2983–2997

Ballarin L, Tonelli C, Guidolin L, Sabbadin A (1999) Purification and characterization of a humoral opsonin, with specificity for D-galactose, in the colonial ascidian *Botryllus schlosseri*. Comp Biochem Physiol 123B:115–123

Bentley DR (1988) Structural superfamilies of the complement system. Exp Clin Immunogenet 5:69–80

Briand JP, Muller S, Van Regenmortal MH (1985) Synthetic peptides as antigens: pitfalls of conjugation methods. J Immunol Methods 78:59–69

Clow LA, Gross PS, Shih C-S, Smith LC (2000) Expression of SpC3, the sea urchin complement component, in response to lipopolysaccharide. Immunogenetics 51:1021–1033

Coombe DR, Ey PL, Jenkin CR (1984) Particle recognition by haemocytes from the colonial ascidian *Botrylloides leachi*: evidence that the *B. leachi* HA-2 agglutinin is opsonic. J Comp Physiol 154:509–521

De Leo G (1992) Ascidian haemocytes and defence reactions. Boll Zool 59:195–214

Feinberg P, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Ann Biochem 132:6–13

Gross PS, Clow LA, Smith LC (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

Isaac L, Isenman DE (1992) Structural requirements for thioester bond formation in human complement component C3: reassessment of the role of thioester bond integrity on the conformation of C3. J Biol Chem 267:10062–10069

Ji X, Azumi K, Sasaki M, 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

Kasahara M (1999) The chromosomal duplication model of the major histocompatibility complex. Immunol Rev 167:17–32

Kelly KL, Cooper EL, Raftos DA (1992) Purification and characterization of a humoral opsonin from the solitary urochordate *Styela clava*. Comp Biochem Physiol 103B:749–753

Kuroda N, Naruse K, Shima A, Nonaka M, Sasaki M (2000) Molecular cloning and linkage analysis of complement C3 and C4 genes of the Japanese medaka fish. Immunogenetics 51:117–128

Lakshma Reddy A, Bryan B, Hildemann WH (1975) Integumentary allograft versus autograft reactions in *Ciona intestinalis*: a

protochordate species of solitary tunicate. Immunogenetics 1:584–590

Lambris JD (1988) The multifunctional role of C3, the third component of complement. Immunol Today 9:387–393

Lambris JD (1990) The third component of complement – chemistry and biology. Springer, Berlin Heidelberg New York

Law SKA, Dodds AW (1997) The internal thioester and the covalent binding properties of the complement proteins C3 and C4. Protein Sci 6:263–274

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

Nagar B, Jones RG, Diefenbach RJ, Isenman DE, Rini JM (1998) X-ray crystal structure of C3d: a C3 fragment and ligand for complement receptor 2. Science 280:1277–1281

Nielsen H, Engelbrecht J, Brunak S, Heijne G von (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1–6

Nonaka M (2001) Evolution of the complement system. Curr Opin Immunol 13:69–73

Nonaka M, Takahashi M (1992) Complete complementary DNA sequence of the third component of complement of lamprey. Implication for the evolution of thioester containing proteins. J Immunol 148:3290–3295

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

Pearce S, Newton RA, Nair SV, Raftos DA (2001) Humoral opsonins of the tunicate, *Pyura stolonifera*. Dev Comp Immunol 25:377–385

Rowley AF, Rhodes CP, Ratcliffe NA (1984) Protochordate leucocytes: a review. Zool J Linn Soc 80:283–295

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sanger F, Coulson AR (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol 94:441–448

Sarrias MR, Franchini S, Canziani G, Argyropoulos E, Moore WT, Sahu A, Lambris JD (2001) Kinetic analysis of the interactions of complement receptor 2 (CR2, CD21) with its ligands C3d, iC3b and the Epstein Barr virus glycoprotein gp350/220. J Immunol 167:1490–1499

Simeone A, Avantaggiato V, Moroni MC, Mavilio F, Arra C, Cotelli F, Nigro V, Acampora D (1995) Retinoic acid induces stage-specific antero-posterior transformation of rostral central nervous system. Mech Dev 51:83–98

Smith LC, Shih C-S, Dachenhausen SG (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 MR, Lambris JD (2000) Complement and innate immunity. Immunopharmacology 49:187–198

Sottrup-Jensen L, Stepanik TM, Kristensen T, Lonblad PB, Jones CM, Wierzbicki DM, Magnusson S, Domdey H, Wetsel RA, Lundwall A, Tack BF, Fey GH (1985) Common evolutionary origin of α_2-macroglobulin and complement components C3 and C4. Proc Natl Acad Sci USA 82:9–13

Sunyer JO, Zarkadis IK, Lambris JD (1998) Complement diversity: a mechanism for generating immune diversity. Immunol Today 19:519–523

Thompson J, Higgins D, Gibson T (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Zarkadis IK, Sarrias MR, Sfyroera G, Sunyer JO, Lambris JD (2001) Cloning and structure of three trout C3 molecules: a plausible explanation for their functional diversity. Dev Comp Immunol 25:11–24