Published in final edited form as:

Fish Shellfish Immunol. 2009 July ; 27(1): 40–49. doi:10.1016/j.fsi.2009.04.001.

Molecular and expression analysis of complement component C5 in the nurse shark (*Ginglymostoma cirratum*) and its predicted functional role

Matthew Graham,

Department of Biological Sciences, Florida International University, University Park, Miami, FL, 33199, USA

Dong-Ho Shin, and

Department of Biological Sciences and Comparative Immunology Institute, Florida International University, Miami, FL, 33199, USA

Sylvia L. Smith*

Department of Biological Sciences and Comparative Immunology Institute, Florida International University, Miami, FL, 33199, USA

Abstract

We present the complete cDNA sequence of shark (*Ginglymostoma cirratum*) pro-C5 and its molecular characterization with a descriptive analysis of the structural elements necessary for its potential functional role as a potent mediator of inflammation (fragment C5a) and initiator molecule (fragment C5b) for the assembly of the membrane attack complex (MAC) upon activation by C5 convertase. In mammals the three complement activation cascades, the classical, alternative and lectin pathways, converge at the activation of C3, a pivotal complement protein. It is, however, the subsequent activation of the next complement component, C5, which is the focal point at which the initiation of the terminal lytic pathway takes place and involves the stepwise assembly of the MAC. The effector cytolytic function of complement occurs with the insertion of MAC into target membranes causing dough-nut like holes and cell leakage. The lytic activity of shark complement results in structurally similar holes in target membranes suggesting the assembly of a shark MAC that likely involves a functional analogue of C5. The composition of shark MAC remains unresolved and to date conclusive evidence has been lacking for shark C5. The gene has not been cloned nor has the serum protein been characterized for any elasmobranch species. This report is the first to confirm the presence of C5 homologue in the shark. GcC5 is remarkably similar to human C5 in overall structure and domain arrangement. The GcC5 cDNA measured 5160-bp with 5' and 3' UTRs of 35bp and 79bp, respectively. Structural analysis of the derived protein sequence predicts a molecule that is a two-chain structure which lacks a thiolester bond and contains a C5 convertase cleavage site indicating that activation will generate two peptides, akin to C5b and C5a. The putative GcC5 molecule also contains the C-terminal C345C/Netrin module that characterizes C3, C4 and C5. Multiple alignment of deduced amino acid sequences show that GcC5 shares more amino acid

*Address correspondence and reprint requests to Sylvia L. Smith, Department of Biological Sciences, Florida International University, 11200 SW 8th Street, Miami, FL, 33199 E-mail addresses: smiths@fiu.edu (S.L. Smith), shindong@fiu.edu (D-H. Shin), Tel.: +1 305 348 3421; Fax: +1 305 348 1083.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

identities/similarities with mammals than that with bony fish. We conclude that at the time of emergence of sharks the elaborate mosaic structure of C5 had already evolved.

Keywords

Shark complement; C5; Gene cloning; Evolution; Phylogenetics; Anaphylatoxin; Membrane attack complex; *Ginglymostoma cirratum*

1. Introduction

Complement, a complex system of soluble and membrane bound proteins, is an integral part of the innate immune system [1–3]. The activation of complement occurs via one or more of three activation pathways, the classical (CP), alternative (AP) and lectin (LP) pathways [4–9]. Each pathway involves the sequential proteolytic cleavage of specific components (present in serum as inactive zymogens) by serine proteases to generate active peptides that facilitate phagocytosis by opsonization of foreign targets and are mediators of the inflammatory response [10–12]. Lysis of target cells is one of the effector functions of mammalian complement. This is achieved by assembly and insertion of the membrane attack complex (MAC), a macromolecular complex composed of complement proteins C5b, C6, C7, C8 and several molecules of C9 [13]. Unlike the activation pathways, the terminal lytic pathway leading to the assembly of MAC involves no proteolytic cleavage of components C6 through C9, but rather activation involves conformation changes following the sequential binding of components. C6 and C7 bind to C5b while the latter is still attached to the C3b moiety of C5 convertase and the binding is thought to occur through sites on the alpha (α)-chain of the C5b fragment [14,15]. The assembly of MAC is initiated following the proteolytic cleavage of C5 that generates two fragments C5b and C5a. C5b sequentially activates C6 through C9 at the site of complement activation and thus is the initiator molecule for complement-mediated cytolysis; while C5a, an anaphylactic peptide, is a mediator of the inflammatory response [13,16–18].

Mammalian C5 is a glycoprotein of 190kDa. The native molecule is composed of two polypeptide chains, an α -chain (115kDa) and a β -chain (75kDa) linked together by a di-sulfide bond [19–21]. The molecule is synthesized as a single chain pro-C5 precursor molecule that undergoes post-translational modification involving glycosylation and specific cleavage at an arginine-rich linker sequence (RPRR) located at the β - α junction to form the two chain structure. Activation of native C5 occurs by proteolytic cleavage of the α -chain at a specific site by C5 convertase (a serine protease enzyme complex) of either the AP or CP to generate C5b (180kDa) and an anaphylactic peptide, C5a (9kDa) [14]. C5a is a small cationic peptide that is a potent anaphylatoxin and chemoattractant which enhances the ability of neutrophils and monocytes to adhere to vessel walls, migrate and phagocytize particles [22,23]. The activity of C5a is controlled by a plasma enzyme, carboxypeptidase N that cleaves the C-terminal arginine of the molecule and thus significantly reduces its activity. C5a desArg retains approximately 1% of its anaphylactic activity [24]. Physiologically, anaphylatoxins are proinflammation peptides mediating smooth muscle contraction, histamine release from mast cells, and increase in vascular permeability. A similar role for this anaphylatoxin has been reported in several teleost fish [25–27].

Due to multiple rounds of tetraploization early in the vertebrate lineage many gene families have arisen [28] including the thiolester-containing protein (TEP) family [29]. C5 is a member of the TEP family which includes two other closely related complement proteins, C3 and C4, and all three are believed to have arisen by gene duplication from α_2 Macroglobulin-like ancestral molecule [30]. Ironically, C5 does not contain a thiolester in its α -chain [31,32].

Previous studies have confirmed the presence of C3 and C4 homologues in shark serum [33, 34], however, direct evidence for the presence of C5 has been lacking although its existence is strongly indicated by complement-derived chemoattractant and spasmogenic activity in zymosan-activated shark serum that is absent from non-activated serum and lost when serum is heated [35]. Furthermore, the migratory response of shark leukocytes to porcine C5a points to the presence of a C5a receptor (CD88) [36–38]. In addition, shark serum complement has been shown to effectively lyse erythrocytes and this ability is lost when the serum is either heated or depleted of key components [34,39–41]. Target cell cytolysis most likely involves the assembly of a MAC-like macromolecular complex. The aim of the present study was to obtain evidence for the existence of C5 homologue in the shark and to determine whether the molecule had structural elements necessary for it to potentially function, upon activation, as an initiating molecule for the assembly of a MAC-like complex and as an anaphylatoxin,

2. Materials and methods

2.1 Materials

The 5'/3' RACE PCR Kit, and PowerScript™ Reverse Transcriptase were obtained from Clontech (Palo Alto, CA, USA). Wizard PlusSV Minipreps DNA Purification System was purchased from Promega (Madison, WI, USA). Restriction enzymes, PCR Supermix High Fidelity, Oligo (dt)_{12–18} primer, TOPO Cloning Kit, and TRIzol reagent were purchased from Invitrogen/Life Technologies (Carlsbad, CA, USA). Big-Dye Terminator Cycle Sequencing Kit (v.3.1) was obtained from Applied Biosystems (Foster City, CA, USA). PCR DIG Probe Synthesis Kit was obtained from Boehringer Mannheim (Indianapolis, IN, USA). Hybond-N⁺ nylon membrane was purchased from Amersham Biosciences (Piscataway, NJ, USA). Lumiphos Plus was purchased from Whatman Biosciences (MA, USA). X-ray film was purchased from Kodak, (New Haven, CT, USA).

2.2 Tissues and isolation of peripheral blood cells

Blood was drawn from the caudal vein of the shark and peripheral blood cells, leukocytes and erythrocytes, were isolated as described previously [42]. Briefly, citrated blood was diluted with shark-RPMI (RPMI with 0.35M urea and 0.25M NaCl) and leukocytes separated on a Ficoll-Paque gradient [43]. Contaminating erythrocytes were removed from harvested leukocytes by hypotonic lysis. The erythrocyte pellet remaining in the gradient tube was collected. Erythrocytes and leukocytes were suspended in shark RPMI at a concentration of 1×10^7 cells/ml, immediately frozen in liquid nitrogen and stored at -80°C .

2.3 First-strand cDNA synthesis

Nurse shark liver was homogenized in TRIzol Reagent (Invitrogen Life Technologies) according to manufacturer's instructions so as to extract total RNA. First-strand cDNA was synthesized using Powerscript II reverse transcriptase (Clontech) and the Oligo(dT)_{12–18} primer (Invitrogen Life technologies) according to the manufacturer's suggested protocol.

2.4 Cloning of the full-length GcC5 cDNA

Degenerate primers (Table 1) covering about 210-bp were designed based on highly conserved sequence in C5 but not in C3, C4 or $\alpha 2\text{M}$ in other taxa as follows; sense primer; 5'-TCNTAYTTNCCNGARACNTGG-3' for SYFPESW (aa 767–773 in human C5) and antisense: 5'-NGGNGCNTGNTGYCAYATNCC for PYSVVRG (aa 830–838 in human C5). The primers were used for PCR amplification of nurse shark total liver cDNA in Perkin Elmer Gene Amp PCR system 2400. The PCR program was as follows: 94°C for 5 min, 40 cycles of 94°C for 30 sec, 45°C for 1 min, and 72°C for 1 min, with a final extension at 72°C for 10 min.

The PCR amplified products were gel-purified using QIAGEN gel extraction kit and cloned into the TOPO TA Cloning® vector. After chemically competent cells (*E. coli* TOP10F') were transformed with ligation, recombinants were identified by blue-white color selection. Vector specific primers (M13 forward and M13reverse, Invitrogen) were used to amplify the plasmid DNA of white colonies by colony PCR.

The clones, whose amplified bands were of expected size, were chosen for sequence analysis. The plasmid DNA was purified from these clones using the Wizard Plus SV Mini-preps kit, according to the manufacturer's instructions. The plasmid DNA (approximately 130ng) then served as the template for a sequencing reaction, carried out in the Core Sequencing facility at Florida International University (Miami, FL, USA).

From the sequence obtained using the degenerate primers, gene specific primers (GSP) were designed for 5' and 3' RACE (rapid amplification of cDNA ends) (Table 1). These primers were used separately in conjunction with the universal primer (Clontech) and the appropriate 3'- and 5'-RACE-ready cDNA, produced using Clontech's Advantage 2 PCR kit. As each consecutive transcript was cloned and sequenced, the overlapping compilation represented the target for full-length amplification. Once the 5' and 3' ends were obtained, primers P1 and P2 (Table 1) were used to amplify the entire cDNA sequence of GcC5. The program was as follows: 94°C for 2 min, 40 cycles of 94°C for 30 sec, 49°C for 30 sec, and 72°C for 7 min, followed by a final extension at 72°C for 13 min. The resulting products were gel-purified and cloned into the TOPO XL Cloning® vector, ideal for longer transcripts. Primers P3-33 (Table 1) were used to determine the sequence of the full-length transcript. The results were truncated as to only allow sequence represented by trustworthy signal, yet they overlapped each other by at least 100 bases.

2.5 cDNA and Phylogenetic analysis

Nucleotide sequences were translated by the BioEdit program [44]. Putative amino acid sequences served as the reference for BLAST (Basic Local Alignment Search Tool) [45] searches, which identified similar Genbank entries. Amino acid sequences corresponding to the gene coding sequences homologous to C5 from various taxa (Fig. 3) were aligned with the putative amino acid sequence of GcC5 clones using the ClustalW software [46] (Gonnet protein weight matrix, default parameters). The alignment was analyzed by the Molecular Evolutionary Genetics Analysis (MEGA 3.1) software, applying the neighbor-joining algorithm to construct a phylogenetic tree (1000 replicates, random seed, Poisson correction, uniform rates).

2.6 Southern blot analysis

Southern blot hybridization was performed to determine the gene copy number of GcC5. The probe was designed based on shark C5 sequence, taking into account the intron/exon boundaries (as determined for the human and mouse orthologues) [20,47,48]. Primers P36 and P37 (Table 1) were designed corresponding to putative exons 30 and 31. To synthesize the probe, a reaction mix containing 5µl 10X PCR buffer, 1.5µl 50mM MgCl₂, 1.6µl 20µM dATP, 1.6µl 20µM dGTP, 1.6µl 20µM dTTP, 1µl each primer, 1ng template (cDNA library from nurse shark spleen), 0.5µl Taq polymerase, 24.2µl sterile water, and 10µl α-³²P-dCTP (100µ Ci) was used in a program set as: 94°C for 4 min, followed by 42 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 30 sec, and a final elongation of 72 °C for 3 min. DNA transfer and blotting was performed on genomic DNA extracted from the erythrocytes of two nurse sharks (separately), digested with *Bam*HI, *Eco*RI, *Hind*III, *Pst*I, and *Sac*I (Roche 10µg DNA per digest) as described by Greenberg et al [49]. Prehybridization and hybridization were performed in the same solution [50% (v/v) formamide/6x SSC/5x Denhardt's solution/0.3% SDS, and denatured salmon sperm DNA at 100 µg/ml (1x SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7.0; 1x Denhardt's solution = 0.02% Ficoll/0.02% polyvinylpyrrolidone/0.02%

bovine serum albumin)] at 42°C, for at least 16 hr for hybridization. The membrane was washed twice in 2x SSC/0.1% SDS for 5 min at room temperature, and then for 20 min in 0.2x SSC/0.1% SDS at 65°C before overnight exposure to film.

2.7 Gene expression analysis of GcC5

The tissue expression of GcC5 was measured semi-quantitatively by RT-PCR. Tissue samples were obtained from a sacrificed nurse shark and frozen in liquid nitrogen and stored at -80°C. Total mRNA from 11 tissues (liver, RBC, WBC, brain, kidney, intestine, ovary, muscle, pancreas, spleen, and heart) was extracted using the TRIzol® reagent kit. First-strand cDNA was synthesized using the PowerScript™ Reverse Transcriptase (RT) according to the manufacturer's protocol. The reaction was terminated by heating the mixture at 70°C for 15min. The primers used in the RT-PCR reaction were the same primers designed for Southern hybridization (P36 and P37). The mix used was 3ul of the first-strand cDNA, 1ul of each primer, and 45ul of PCR SuperMix High Fidelity, amplified using the program: 94°C for 2 min, then 37 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 30 sec, followed by a final elongation of 72°C for 3 min. The expression of β -actin was the positive control using primers P34 and P35. The amplification of the positive control was carried out using the program: 94°C for 1 min, then 31 cycles of 94°C for 30 sec, 51°C for 30 sec, and 72°C for 55 sec, followed by a final elongation period at 72°C for 5 min.

3. Results

3.1. Identification and sequencing of full-length GcC5 cDNA

Degenerate PCR amplification produced a sequence of 218bp in length, which was then gel purified and cloned into the TOPO TA vector. Approximately 60 clones were chosen for sequence analysis. One sequence was identified as having similarity to C5, and served as the template for the design of gene specific primers (GSP) used in 5' and 3' RACE. The 5' and 3' RACE PCR amplifications using the GSPs mentioned above yielded fragments of approximately 400bp and 1800bp in length, respectively. A sufficient number of PCR amplified products were used in the cloning of each amplicon, and subsequently, plasmid extraction and sequencing. Once the entire GcC5 cDNA was compiled by overlapping clones, the entire sequence was amplified as one transcript so as to rule out the possibility of a chimeric GcC5 gene. GcC5 is 5160bp in length with an open reading frame of 5046bp, with 5' and 3' UTRs of 35bp and 79bp, respectively (Fig. 1). GcC5 has start and stop codons (located at positions 36 and 5079, respectively), and a polyadenylation signal (AATAAA) located 21 nucleotides upstream from the polyadenylation site. The β - α cut site (RPKR) located at aa positions 674–677 in GcC5, although similar to that of carp and trout C5, differs from that of human (RPRR) C5.

3.2 Multiple alignment

Multiple alignment of the putative amino acid sequence of nurse shark GcC5 and thiolester family members from other taxa (Fig. 2) allowed for the characterization of functional domains. Alignments were performed using the ClustalW software [46], and shaded residues represent a 75% level of similarity (gray) and identity (black). The beginning of the β -chain, β - α cut site, the beginning of the α -chain and the C5 convertase cleavage site are indicated by arrows. The GcC5a anaphylatoxin region shares many of the same characteristics as human C5a such as the six conserved cysteines that provide molecular stability. The C-terminal effector region of GcC5a shares 5 of 9 residues with human C5a, including the terminal arginine which is essential for optimal function [50].

The region of thiolester bond formation (in C3 and C4) is shown in bold italics. The residues responsible for forming the thiolester bond, cysteine and glutamine [51–53], are lacking in

GcC5. The final 145 residues resemble a region denoted C345C-Netrin domain [54–56]. Within this region, a putative N-linked glycosylation site (located at N1582) sits roughly 50 residues upstream from a similar glycosylation site in human C5. Other potential N-linked glycosylation sites are denoted with diamonds. The cleavage site between the signal peptide and the β -chain was determined using the SignalIP 3.0 server [57,58].

Shark GcC5 contains three N-linked glycosylation sites (N⁹¹⁶, N¹⁵¹³, and N¹⁵⁸²) on the α -chain in contrast to human C5 which has 4 sites. GcC5 does not contain an N-linked glycosylation site within its C5a region in contrast to human C5 which has one site and is a glycosylated peptide [59]. In this respect GcC5a resembles porcine C5a which also lacks an N-linked glycosylation site. Teleost C5 molecules, such as carp C5-1 and C5-2 [61], contain two and three N-linked glycosylation sites, respectively. In carp C5-1 both sites are on the β chain while in C5-2 only one is on the β chain and the remaining two are present on the α chain. Similar to shark GcC5a sequence, neither of the carp C5 molecules contains an N-linked glycosylation site within the C5a region.

3.3 Phylogenetic analysis

The evolutionary relationship of GcC5 was examined with respect to the members of the thiolester-containing protein family by constructing a phylogenetic tree. The translated coding sequence of GcC5 was placed as a sister to the bony fish C5 and mammalian C5 clades, with the most similarity to rat C5. GcC5 also appears to be more closely related to C3 than C4 (Fig 3). Phylogenetic analysis shows that C5 of cartilaginous fish has evolved from a common ancestor of mammals and bony fish and that GcC5 is closer to mammalian C5 than bony fish C5. It further indicates that the emergence of complement component C5 followed that of C3.

3.4 Southern Blot analysis

Southern hybridization was used to predict the number of copies of the GcC5 gene. The cDNA probe expected to span two exons (based on the exon-intron structure of C5 in the human and the mouse) [20,47,48] and lacking cut sites for the five restriction enzymes utilized, was used to distinguish C5 within the genomic DNA from two nurse sharks. Two bands are seen in four of the five digests, the lone exception (*EcoRI*) with a single band that migrates higher than all the others (Figure 4). This pattern suggests at most two loci for the C5 gene in shark. Duplicated loci could be close enough that *EcoRI* does not cut in the intronic region probed. However we have no sequence data to suggest greater than two alleles from an individual that would corroborate more than one locus. Therefore, the gene may be single copy and allelic polymorphisms affecting restriction sites in the intron could be responsible for the doublets in some digests.

3.5 Expression analysis of GcC5 gene

The same primers used to make the Southern probe were used in an RT-PCR amplification from nine tissues and peripheral blood cells (red and white blood cells) to examine the expression of GcC5. Figure 5 shows the band intensities of each tissue sample relative to the β -actin expression from the same sample. As is consistent with the literature, the liver appears to be the primary site of synthesis for C5 [2], followed by the ovary and pancreas with high levels of expression. With a somewhat intermediate expression, as compared to the other tissues, was the kidney, intestine, RBC, WBC, spleen, and heart. The lowest, but still detectable level of expression was in the brain and muscle. From results shown in Figure 5 we can infer that C5 may be constitutively expressed in all the sample tissues of the shark albeit there are differences in the level of gene expression.

4. Discussion

The existence of a C5-like protein in an elasmobranch has not been reported. The aim of the present study was to identify and characterize the gene encoding the human C5 homologue in the shark and to determine whether the molecule had structural elements necessary for it to function as a lead-in molecule for MAC assembly and as a source of anaphylatoxin. A full-length cDNA encoding putative shark C5 protein (GcC5) has been isolated from shark liver and sequence analysis of GcC5 cDNA confirms it to be a homologue of C5. Derived protein sequence of GcC5 shows a high degree of homology to human C5 (63% similarity), less to bony fish C5 (57 – 59%). Structural analysis reveals that the domain structure is similar to C5 molecules described for human and other vertebrates [20,48,60–65].

The GcC5 cDNA measured 5160bp with 5' and 3' UTRs of 35bp and 79bp, respectively. It encodes a pro-GcC5 protein of 1681 amino acid residues with a signal peptide of 18 residues. The deduced amino acid sequence contains a RPKR linkage sequence identifying it as the potential β - α cut site yielding a two-chain molecule. In addition a second cut site considered essential for C5 convertase cleavage is present that would generate two fragments similar to C5b and C5a. Amino acid sequence analysis of the α -chain of GcC5 does not, however, show a cleavage site for factor I cleavage indicating that the molecule will not be susceptible to factor I inactivation. Susceptibility to factor I is a feature of mammalian C3 and C4 but not C5 since it too lacks the factor I cleavage site. The characteristic sequence triplet Asn-X-Ser/Thr that is diagnostic of asparaginyl-linked carbohydrates [66] has been identified at positions 916, 1513 and 1582 on the α '-chain. We view these asparagine residues as potential oligosaccharide attachment sites, however, while it is highly likely that GcC5 is a glycosylated protein evidence that one or more sites are glycosylated in the shark remains to be obtained.

Mammalian anaphylatoxins, C3a, C4a and C5a are highly cationic peptides characterized by cysteine residues that form three intra-chain disulfide bonds imparting significant chemical and physical stability to the peptides. Analysis of the GcC5a peptide sequence reveals the molecule has the necessary cysteine residues to form multiple intra-chain disulfide linkages (the disulfide core spans residues 700 through 735) which confers molecular conformational stability. In addition, GcC5a contains a free cysteine at position 714, however, it lacks a glycosylation site indicating that unlike human C5a GcC5a is non-glycosylated. The C-terminal sequence MQLGR of human C5a is the effector region of the molecule. A binding structural region that acts independent of its effector region (the disulphide core) has been identified [54]. The C-terminal putative effector region of the GcC5a anaphylatoxin sequence contains the sequence LTLGR which contains the critical terminal arginine residue that is considered essential for anaphylatoxic activity. The C-terminal sequence of GcC5a is distinct from the C-terminal sequences, SQMTLAR and TQMTLAR, of GcC3a-1 and GcC3a-2, respectively, and KVDSIAR of TrscC4a of the houndshark (*Triakis scyllia*). The former two peptide sequences are of the anaphylatoxin fragments of two C3 genes in the shark, GcC3-1 and GcC3-2 (Smith, unpublished). Taken together the sequence data shows that GcC5 is a TEP molecule distinct from other shark TEPs (GcC3s and TrscC4).

Complement TEPs, C3, C4 and C5 have a unique feature in common which distinguishes them from other members of the TEP superfamily of proteins. Common to all three molecules is a ~150-residue-long C-terminal segment which contains the characteristic C345C module [54, 56]. The amino acid residues within the C345C module of C5 are considered important for the activation of C5 by C5 convertase [1,55]. Unlike C3, in which C345C is not involved with direct binding of the molecule with other complement proteins, C5-C345C has been shown to play an important role in binding activities of the molecule, specifically with C6 and C7 [67]. Activation of C5 involves specific cleavage of an arginyl-leucine peptide bond in the α -chain with the release of C5b, which, once formed is labile and decays unless it quickly associates

with C6 [68]. C5b undergoes a conformational change involving expression of hydrophobic residues such that it acquires a metastable binding site for C6 followed by C7 and it has been suggested that the binding occurs between the C-terminal C345C module of C5b and FIMs (Factor I modules, also referred to as FIMAC modules) of C6/C7, present at the C-terminal of both proteins but absent from C8 and C9 [54,56]. The bond between C5-C345C domain of C5 and C7-FIMs is considered essential for interaction between C5bC6 and C7 to occur. C5b6-7 is able to integrate into the outer face of phospholipid bilayer membranes through hydrophilic to amphiphilic transformation and the addition of C8 and C9(n), results in the formation of trans-membrane channel complex (MAC) on target membrane. Whether the corresponding module in shark, GcC5-C345C, plays a similar role is uncertain since the existence of C6 and C7 homologues in shark remains speculative. As human C5 can bind directly to C7 through the latter's FIMs in the absence of C6, gives rise to the speculation that in the shark it might not be necessary to have both proteins present in order to assemble a functional MAC. Thus, it is reasonable to consider that the assembly of shark MAC in addition to involving GcC5b, and terminal components C8 and C9 (functional analogues of both are present in shark serum) [40,69] will involve at least one additional protein, a C6- or C7-like protein. Alternatively, it is also possible that in the shark, in the absence of functional analogues of C6 and C7, GcC5b could conceivably react directly with C8 and serve as a potential anchor for shark C8. Earlier studies on mammalian complement have shown this direct interaction between C5 and C8 to occur. The C5b site associates with the beta subunit of C8 which appears to use the same site for both reversible and non-reversible binding to C5 [67,70,71]. While the presence of C8 β subunit homologue has not been confirmed, studies have shown the existence of C8 α homologue in the shark (Aybar and Smith, unpublished).

In mammals the liver is the main site of C5 synthesis, however, it is also synthesized by macrophages [72]. In the shark constitutive expression of GcC5 was seen in most tissues examined including blood leukocytes with the highest level being in the liver, GcC5 expression pattern is similar to that noted in other fish species. Southern blot hybridization analysis suggests the presence of a single copy GcC5 gene. In teleosts several copies of C5 genes have been reported [61]. Phylogenetic studies place GcC5 within the C5 clade, basal to mammalian homologues.

In summary, our studies show that GcC5 is synthesized as a single chain precursor molecule of 1681 amino acid residues with an 18 residue leader peptide and an arginine-rich linker sequence (RPKR) located at the β - α junction, predicting the native molecule to be a two chain structure. The presence of a cut site for C5 convertase indicates that cleavage will result in the generation of C5b and C5a fragments. Since GcC5 lacks the sequence CGEQ that denotes thiolester bond in its α -prime chain the GcC5b fragment will, therefore, lack the ability to covalently link to target surfaces, a feature seen with opsonic peptides C3b and C4b. Structural analysis also reveals three potential N-linked glycosylation sites suggesting the native molecule is probably a glycoprotein. The putative C5a peptide is 77 amino acid residues in length. The conserved cysteine residues suggest preservation of a preferred folding pattern that will provide molecular stability required for maximal functional activity. The last five residues at the C-terminus suggest a biologically active anaphylatoxic peptide. The overall structure and domain arrangement of GcC5 shows that most if not all the structural features are present that are required for it to serve a functional purpose similar to its mammalian homologue. This study is the first to establish the presence of complement component C5 homologue in a lower chordate, i.e., an elasmobranch (cartilaginous fish). Structural analysis provides compelling evidence for its predicted functional role in the assembly of a membrane attack complex (essential for complement-derived cytolytic activity) and as an inflammatory mediator. The study contributes significantly to furthering our understanding of evolution of the TEP family of proteins and from the data we also conclude that at the time of the emergence of sharks the

elaborate mosaic structure of C5 had already evolved. Studies to define the functional role of GcC5 are underway.

Abbreviations and definitions

AP	alternative pathway
C	complement
C1,C2,C9	complement components 1 through 9
CP	classical pathway
AP	alternative pathways
LP	lectin pathway
EA	antibody-sensitized erythrocytes
EAC1-7	EA reacted sequentially with the first seven mammalian C components

References

1. Müller-Eberhard HJ. Molecular organization and function of the complement system. *Annu Review of Biochemistry* 1988;57:321.
2. Volanakis J. Overview of the complement system: The human complement system in health and disease. *Volanakis, J cop* 1998;9–31.
3. Lambris J, Reid K, Volanakis J. The evolution, structure, biology and pathophysiology of complement. *Immunology Today* 1999;20:207–211. [PubMed: 10322298]
4. Hugh-Jones, N. The classical pathway. In: Ross, GD., editor. *Immunobiology of the complement system*. New York: Academic Press; 1986. p. 21-44.
5. Pangburn MK, Müller-Eberhard HJ. The alternative pathway of complement. *Springer Semin Immunopath* 1984;7:63–192.
6. Ohta M, Okada M, Yamashina I, Kawasaki T. The mechanisms of carbohydrate-mediated complement activation by the serum mannose-binding protein. *J Biol Chem* 1990;265:1980–1984. [PubMed: 2298734]
7. Sato T, Endo Y, Matsushita M, Fujita T. Molecular characterization of a novel serine protease involved in activation of the complement system by mannose-binding protein. *Int Immunol* 1994;6:665–669. [PubMed: 8018603]
8. Law, SKA.; Reid, KBM. Complement. In: Male, D., editor. *Complement*. Vol. 2. IRL Press; Oxford: 1995.
9. Whaley, Schwaebler W. Complement and complement deficiencies. *Semin Liver Dis* 1997;17:297–310. [PubMed: 9408965]
10. Tomlinson S. Complement defense mechanisms. *Curr Opin Immunol* 1993;5:83–89. [PubMed: 8452679]
11. Matsushita M, Fujita T. Activation of the classical complement pathway by mannose binding protein with a novel C1s-like serine protease. *J Exp Med* 1992;176:1497–1502. [PubMed: 1460414]

12. Matsushita M, Fujita T. Cleavage of the third component of complement (C3) by mannose-binding protein-associated serine protease (MASP) with subsequent complement activation. *Immunobiol* 1995;194:443–448.
13. Podack, ER. Assembly and structure of the membrane attack complex (MAC). In: Podack, editor. *Cytolytic lymphocytes and complement: effectors of the immune system*. Vol. 1. Boca Raton: CRC Press; 1988. p. 173–184.
14. Vogt W, Schmidt E, Von Buttlar B, Dieminger LA. A new function of the activated third component of complement: binding to C5, an essential step for C5 activation. *Immunology* 1978;34:29–30. [PubMed: 624565]
15. Thai CT, Ogata RT. Expression and characterization of the C345C/NTR domains of complement components C3 and C5. *J Immunol* 2003;171:6565–6573. [PubMed: 14662858]
16. Podack, ER. Assembly and functions of the terminal components. In: Ross, GD., editor. *Immunobiology of the complement system*. New York: Academic Press; 1986. p. 115–137.
17. Hugli TE, Müller-Eberhard. Anaphylatoxins: C3a and C5a. *Adv Immunol* 1978;25:1–55. [PubMed: 358800]
18. Hugli TE. Structure and function of the anaphylatoxins. *Springer Semin Immunopathol* 1984;7:193–219. [PubMed: 6387982]
19. Fernandez HN, Hugli TE. Chemical evidence for genetic ancestry of complement components C3 and C4. *J Biol Chemistry* 1977;252:1826–1828.
20. Tack BF, Morris SC, Prahl JW. Fifth component of human complement: purification from plasma and polypeptide chain structure. *Biochemistry* 1979;18:1490–1497. [PubMed: 106884]
21. Wetsel RA, Lemons RS, Le Beau MM, Barnum SR, Noack D, Tack BF. Molecular analysis of human complement component C5: localization of the structural gene to chromosome 9. *Biochemistry* 1988;27:1474–1482. [PubMed: 3365401]
22. Hugli T. Biochemistry and biology of anaphylatoxin. *Complement* 1986;3:111–127. [PubMed: 3542363]
23. Gerard C, Gerard N. C5a anaphylatoxin and its seven transmembrane segment receptor. *Annual Review of Immunology* 1994;12:775–808.
24. Gerard C, Hugli TE. C5a: a mediator of chemotaxis and cellular release reactions. *Kroc Found Ser* 1981;14:147–160. [PubMed: 6975805]
25. Sunyer JO, Tort L. The complement system of the teleost fish *Sparus aurata*. *Ann. N.Y. Acad. Sci* 1994;712:371.
26. Sunyer JO, Boshra H, Li J. Evolution of anaphylatoxins their diversity and novel roles in innate immunity: insights from the study of fish complement. *Vet Immunol Immunopath* 2005;108:77–89.
27. Kato Y, Nakao M, Shimizu M, Wariishi H, Yano T. Purification and functional assessment of C3a, C4a and C5a of the common carp (*C. caprio*) complement. *Dev Comp Immunol* 2004;28:901–910. [PubMed: 15183031]
28. Kasahara M, Nakaya J, Satta Y, Takahata N. Chromosomal duplication and the emergence of the adaptive immune system. *Trends Genet* 1997;13:90–92. [PubMed: 9066265]
29. Law SKA, Dodds AW. The internal thioester and the covalent binding properties of the complement proteins C3 and C4. *Protein Sci* 1997;6:263. [PubMed: 9041627]
30. Dodds A, Law S. The phylogeny and evolution of the thiolester bond-containing proteins C3, C4 and a2-macroglobulin. *Imm Reviews* 1998;166:15–26.
31. Lambris, JD.; Sahu, A.; Wetsel, R. The chemistry and biology of C3, C4 and C5. In: Volanakis, JE.; Frank, M., editors. *The human complement system in Health and Disease*. Marcel Dekker; New York: 1998. p. p83–118.
32. Sottrup-Jensen L, Stepanik TM, Kristensen T, Lonblad PB, Jones CM, Wierzbicki DM, Magnusson S, Domdey H, Wetsel RA, Lundwall A, et al. Common evolutionary origin of alpha 2-macroglobulin and complement components C3 and C4. *Proc Natl Acad Sci USA* 1985;82:9–13. [PubMed: 2578664]
33. Dodds AW, Smith SL, Levine RP, Willis AC. Isolation and initial characterisation of complement components C3 and C4 of the nurse shark and the channel catfish. *Developmental and Comparative Immunology* 1998;22:207–16. [PubMed: 9639090]

34. Smith S. Shark complement: an assessment. *Immunological Reviews* 1998;166:67–78. [PubMed: 9914903]
35. Smith S, Riesgo M, Obenauf S, Woody C. anaphylactic and chemotactic response of mammalian cells to zymosan activated shark serum. *Fish and Shellfish Immunology* 1997;7:503–514.
36. Obenauf SD, Smith SH. Chemotaxis of nurse shark leukocytes. *Dev Comp Immunol* 1985;9:221–30. [PubMed: 4018340]
37. Hyder Smith S, Obenauf SD. Fine structure of shark leukocytes during chemotactic migration. *Tissue Cell* 1989;21:47–58. [PubMed: 18620254]
38. Obenauf SD, Hyder Smith S. Migratory response of nurse shark leucocytes to activated mammalian sera and porcine C5a. *Fish and Shellfish Immunology* 1992;2:173–181.
39. Legler DW, Evans EE. Comparative Immunology: hemolytic complement in elasmobranchs. *Proceedings of the Society for Experimental Biology and Medicine* 1967;124:30–4. [PubMed: 4959767]
40. Jensen JA, Festa E, Smith D, Cayer M. The complement system of the nurse shark: hemolytic and comparative characteristics. *Science* 1981;214:566–569. [PubMed: 7291995]
41. Culbreath L, Smith SL, Obenauf SD. Alternative complement pathway activity in nurse shark serum. *Am Zool* 1991;31:131A.
42. Shin D-H, Webb B, Nakao M, Smith SL. Molecular cloning, structural analysis and expression of complement component Bf/C2 genes in the nurse shark, *Ginglymostoma cirratum*. *Dev. Comp. Immunol* 2007;31:1168–1182.
43. Wittman G. Separation of porcine blood cells by means of Ficoll-Paque. *Zbl Vet Med B* 1980;27:253–256.
44. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 1999;41:95–98.
45. Altschul S, Gish W, Miller W, Myers E, Lipman D. Basic Local Alignment Search Tool. *Journal of Molecular Biology* 1990;215:403–410. [PubMed: 2231712]
46. Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 1994;22:4673–4680. [PubMed: 7984417]
47. Nilsson UR, Mandle RJ, McConnell-Mapes JA. Human C3 and C5: subunit structure and modification by trypsin and C42. *J Immunol* 1975;114:815–822. [PubMed: 1112981]
48. Wetsel RA, Ogata RT, Tack BF. Primary structure of the fifth component of murine complement. *Biochemistry* 1987;26:737–743. [PubMed: 2436653]
49. Greenberg AS, Steiner L, Flajnik MF. Isolation of a shark immunoglobulin light chain cDNA clone encoding a protein resembling mammalian kappa light chain: implication for the evolution of light chains. *Proc Nat'l Acad Sci USA* 1993;90:10603–10607.
50. Huber-Lang MS, Sarma JV, McGuire SR, Lu KT, Padgaonkar VA, Younkin EM, Guo RF, Weber CH, Zuiderweg ER, Zetoune FS, Ward PA. Structure-function relationships of human C5a and C5aR. *Journal of Immunology* 2003;170:6115–6124.
51. Hostetter MK, Thomas ML, Rosen FS, Tack BF. Binding of C3b proceeds by a transesterification reaction at the thiolester site. *Nature* 1982;298:72–75. [PubMed: 6806664]
52. Tack BF, Harrison RA, Janatova J, Thomas ML, Prahl JW. Evidence for presence of an internal thiolester bond in third component of human complement. *Proc Natl Acad Sci U S A* 1980;77:5764–5768. [PubMed: 6934510]
53. Thomas ML, Janatova J, Gray WR, Tack BF. Third component of human complement: Localization of the internal thiolester bond. *Proc Natl Acad Sci U S A* 1982;79:1054–1058. [PubMed: 6175959]
54. Low PJ, Ai R, Ogata RT. Active sites in complement components C5 and C3 identified by proximity to indels in the C3/4/5 protein family. *J Immunology* 1999;162:6580. [PubMed: 10352274]
55. Sandoval A, Ai R, Ostresh JM, Ogata RT. Distal recognition site for classical pathway convertase located in the C345C/netrin module of complement component C5. *J Immunol* 2000;165:1066. [PubMed: 10878385]

56. Thai C-T, Ogata RT. Expression and Characterization of the C345C/NTR Domains of Complement Components C3 and C5. *J Immunol* 2003;171:6565–6573. [PubMed: 14662858]
57. Nielsen, H.; Krogh, A. Prediction of signal peptides and signal anchors by a hidden Markov model. *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology (ISMB 6)*; Menlo Park, California: AAAI Press; 1998. p. 122-130.
58. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004;340:783–795. [PubMed: 15223320]
59. Fernandez HN, Hugli TE. Primary structural analysis of the polypeptide portion of human C5a anaphylatoxin. Polypeptide sequence determination and assignment of the oligosaccharide attachment site in C5a. *J Biol Chem* 1978;253:6955–6964. [PubMed: 690134]
60. Franchini S, Zarkadis IK, Sfyroera G, Sahu A, Moore WT, Mastellos D, LaPatra SE, Lambris JD. Cloning and purification of the rainbow trout fifth component of complement (C5). *Dev Comp Immunol* 2001;25:419–430. [PubMed: 11356221]
61. Kato Y, Nakao M, Mutsuro J, Zarkadis IK, Yano T. The complement component C5 of the common carp (*Cyprinus carpio*): cDNA cloning of two distinct isotypes that differ in a functional site. *Immunogenetics* 2003;54:807–15. [PubMed: 12618914]Epub 2003 Feb 7
62. Kumar KG, Ponsuksili S, Schellander K, Wimmers K. Molecular cloning and sequencing of porcine C5 gene and its association with immunological traits. *Immunogenetics* 2004 Mar;55:811–7. [PubMed: 14770272]Epub 2004 Feb 10
63. Lambris J, Sahu A, Wetsel R. The chemistry and biology of C3, C4, and C5: The human complement system in health and disease. *Volanakis J* 1998;p83–117.
64. Ooi YM, Colten HR. Biosynthesis and post-synthetic modification of a precursor (Pro-C5) of the fifth component of mouse complement (C5). *J Immunol* 1979;123:2494.
65. Woods E, Ogden RC, Colten HR, Tack BF. Isolation and sequence of a cDNA clone encoding the fifth complement component. *J Biological Chemistry* 1985;260:2108.
66. Clamp, JR. *The Plasma Proteins*. Putnam, F., editor. Vol. 2. Academic Press; New York: p. 163-211.
67. DiScipio RG. Formation and structure of the C5b-7 complex of the lytic pathway of complement. *J Biological Chemistry* 1992;267:17087.
68. Cooper NR, Müller-Eberhard HJ. The reaction mechanism of human C5 in immune hemolysis. *J Exp Med* 1970;132:775–793. [PubMed: 5508377]
69. Jensen JA, Fuller L, Iglasias E. The terminal components of the nurse shark C-system. *J Immunol* 1973;111:306–307.
70. Stewart JL, Kolb WP, Sodetz JM. Evidence that C5b recognizes and mediates C8 incorporation into the cytolytic complex of complement. *J Immunology* 1987;139:1960–1964. [PubMed: 3624872]
71. Kolb WP, Haxby JA, Arroyave CM, Müller-Eberhard HJ. The membrane attack mechanism of complement: reversible interactions among the five native components in free solution. *J Exp Med* 1973;138:428. [PubMed: 4737061]
72. Hetland G, Johnson E, Falk J, Eskeland T. Synthesis of complement components C5, C6, C7, C8 and C9 in vitro by human monocytes and assembly of the terminal complement complex. *Scand J Immunol* 1986;24:421–428.

(in C3 and C4) are shown in bold and italics. Potential N-linked glycosylation sites are denoted with diamonds. The cleavage site between the signal peptide and the β -chain was determined using the SignalIP 3.0 server [57,58]. The C345C-Netrin region is shown both underlined and italicized, with the conserved cysteines indicated with plus signs.

Fig. 3.

Phylogenetic analysis of GcC5 and other thiolester family members. ClustalW aligned numbers to the right of the tree represent the respective percentage of identity and similarity for each protein sequence in pair-wise comparison to GcC5 sequences. The tree was built by the neighbor-joining algorithm using the MEGA 3.1 program, and the numbers within the tree indicate bootstrap support for each partition. The accession numbers for the amino-acid sequences used in tree construction are as follows: carp α 2M, **BAA85038**; carp C3, **BAA36621**; carp C4, **BAB03284**; carp C5-1, **BAC23057**; carp C5-2, **BAC23058**; chicken C3, **NP_990736**; cobra C3, **AAA49385**; cow C4, **p01030**; CVF, **Q91132**; frog C4, **NP_001080934**; fruit fly TEP-1, **NP_523578**; fruit fly TEP-2, **NP_523506**; guinea pig C3, **P12387**; hagfish C3, **P98094**; horseshoe crab α 2M, **BAA19844**; houndshark C4, **BAC82347**; human α 2M, **EAW88590**; human C3, **AAI50180**; human C4, **P0C0L4**; human C5, **P01031**; lamprey C3, **AAR13241**; mosquito α 2M, **XP_001660377**; mosquito TEP, **AAG00600**; mouse α 2M, **AAH72642**; mouse C3, **NP_033908**; mouse C4, **P01029**; mouse C5, **P06684**; pig C3, **NP_999174**; pig C5, **NP_001001646**; prawn α 2M, **BAC99073**; rat C3, **NP_058690**; rat C4, **P08649**; rat C5, **XP_001079130**; shark C5 **EU797190**; snail TEP, **BAE44110**; soft coral C3-like, **AAN86548**; trout C3, **AAB05029**; trout C5, **AAK82852**; tunicate α 2M, **NP_001027688**; urchin C3, **NP_999686**; zebrafish α 2M, **XP_001332356**; zebrafish C4, **XP_694622**.

Fig. 4.

Southern hybridization analysis of genomic DNA from two nurse sharks probed with nurse shark GcC5. Ten micrograms of shark genomic DNA was used for each digestion. Restriction endonucleases are marked at the top as follows: B; *Bam*HI, E; *Eco*RI, H; *Hind*III, P; *Pst*I, and S; *Sac*I. A standard DNA mass marker (λ /*Hind*-III digests) is shown in left with base-pair (bp).

Fig. 5.

RT-PCR analysis of GcC5 mRNA expression in nurse shark tissues and peripheral blood cells. The letters on the bottom indicate names of tissue investigated for RT-PCR in this experiment: K= kidney, S= spleen, B= brain, L= liver, I= intestine, O= ovary, M= muscle, H= heart, P= pancreas, R= erythrocytes, W= leukocytes. β -actin expression in each sample was added for normalization. The size of amplified product was 147 bp for GcC5 and 682 bp for β -actin. Amplification was not seen in controls in the absence of template or enzyme (result not shown).

Table 1

Primers used for sequence analysis, synthesizing DIG-probes, and RT-PCR analysis of tissue expression.

Primer name	Sequence (5'-3')	Position in sequence	Strand	Utility
P1	ACTGTGTACTGAGGATCTG	44–62	Sense	Full length transcript
P2	CATATCTGATAACATCCATAGC	5135–5114	Anti-sense	Full length transcript
P3	GGTATCAACGCAGAGTACG	8–26	Sense	Sequencing
P4	CTGACAAACCCGTTTATACCC	454–474	Sense	Sequencing
P5	GACTTCAAGATTCCTCTAATCC	624–646	Sense	Sequencing
P6	ACAATTACATCAGCATACG	840–858	Sense	Sequencing
P7	CTTGGTAGCAACTCCACTC	1124–1142	Sense	Sequencing
P8	GTCGCAAGTAGTCCCTACTTG	1524–1544	Sense	Sequencing
P9	CGCTATCTTCCATTGACAC	1867–1885	Sense	Sequencing
P10	GATTTGAAGGTGATGGCTG	2097–2115	Sense	Sequencing
P11	GATTGAAATTCAGAGAGCTCAGA	2402–2425	Sense	Sequencing
P12	TCAGCACTTCCATTACATTAC	2754–2776	Sense	Sequencing
P13	AATCTACCTCGTGGCTCTGC	3087–3106	Sense	Sequencing
P14	TCTCCAATTATGTGCCTTCG	3448–3467	Sense	Sequencing
P15	TAACAGCACGAATAGTAGA	3805–3823	Sense	Sequencing
P16	ATTAGCGACCCAGTGAAAC	4179–4198	Sense	Sequencing
P17	GCGTTGGATTCCGTGTTTATG	4495–4515	Sense	Sequencing
P18	TCATCAAGTACAACGCCAC	4798–4816	Sense	Sequencing
P19	TGTGCTTGATGGCCTTAATC	4970–4951	Anti-sense	Sequencing
P20	CATTTCATTCATCTCCAC	4657–4639	Anti-sense	Sequencing
P21	CTGGGTGCGCTAATGACATA	4192–4173	Anti-sense	Sequencing
P22	TGCAACACAGCCAGAAGAGC	3856–3837	Anti-sense	Sequencing
P23	AGTGTGCTCTTCACATTCTC	3640–3621	Anti-sense	Sequencing
P24	TCAGCAGAGCCACGAGGTAG	3109–3090	Anti-sense	Sequencing
P25	GTGAATGTAAATGGAAGTGCTGA	2776–2754	Anti-sense	Sequencing
P26	ACCTCTCACACGAATAAGG	2588–2568	Anti-sense	Sequencing
P27	TCGTCCCAATGTCAGGTCAC	2324–2305	Anti-sense	Sequencing
P28	TCCAGCTCCACATCCAAG	1976–1959	Anti-sense	Sequencing
P29	GCCAACTGGACGGTCTTTG	1625–1607	Anti-sense	Sequencing
P30	AACAGTGGTATCGCCTACTGG	1223–1203	Anti-sense	Sequencing
P31	CCCTCTGGATCCTGAAAAGTG	568–548	Anti-sense	Sequencing
P32	ACGTATGCTGATGTAATTG	859–841	Anti-sense	Sequencing
P33	TATTGTAACCTCCCTCTG	284–265	Anti-sense	Sequencing
P34	CTGCCATGTATGTTGCCATC		Sense	β-ACTIN
P34	ATCCACATCTGCTGGAAGGT		Anti-sense	β-ACTIN
P36	TATGGAAATCCCATTGTTC	3873–3891	Sense	RT-PCR/Southern
P37	CTGAATGTGGCCAAGTC	4049–4033	Anti-sense	RT-PCR/Southern