Published in final edited form as:

Fish Shellfish Immunol. 2009 September ; 27(3): 397–406. doi:10.1016/j.fsi.2009.05.020.

Molecular characterization of the alpha subunit of complement component C8 (GcC8 α) in the nurse shark (*Ginglymostoma cirratum*)

Lydia Aybar¹, Dong-Ho Shin^{1,2}, and Sylvia L. Smith^{1,2,*}

¹ Department of Biological Sciences, Florida International University, Miami, Florida, 33199

² Comparative Immunology Institute, Florida International University, Miami, Florida, 33199

Abstract

Target cell lysis by complement is achieved by the assembly and insertion of the membrane attack complex (MAC) composed of glycoproteins C5b through C9. The lytic activity of shark complement involves functional analogues of mammalian C8 and C9. Mammalian C8 is composed of α , β , and γ subunits. The subunit structure of shark C8 is not known. This report describes a 2341 nucleotide sequence that translates into a polypeptide of 589 amino acid residues, orthologue to mammalian C8 α and has the same modular architecture with conserved cysteines forming the peptide bond backbone. The C8 γ -binding cysteine is conserved in the perforin-like domain. Hydrophobicity profile indicates the presence of hydrophobic residues essential for membrane insertion. It shares 41.1% and 47.4 % identity with human and *Xenopus* C8 α respectively. Southern blot analysis showed GcC8 α exists as a single copy gene expressed in most tissues except the spleen with the liver being the main site of synthesis. Phylogenetic analysis places it in a clade with C8 α orthologs and as a sister taxa to the *Xenopus*.

Keywords

shark complement; membrane attack complex (MAC); C8 α ; eighth complement component C8 alpha; lytic pathway; innate immunity; elasmobranch; *Ginglymostoma cirratum*

1. Introduction

Mammalian complement, an integral component of innate immunity, is a complex system of soluble and cell-bound proteins which can be activated via one or more of three pathways, the classical, alternative or lectin [1,2]. Each activation cascade leads to the terminal lytic or membrane attack pathway resulting in disruption of target membranes. Complement-mediated lysis of target cells involves the formation and insertion of the membrane attack complex (MAC) into the target membrane to form trans-membrane pore-like structures that make cells leaky [3–7]. MAC formation is initiated by the generation of C5b (activation peptide released from C5 by C5 convertase cleavage) followed by the non-enzymatic sequential assembly of complement proteins C6, C7, C8 (α , β , and γ subunits), and several molecules of C9. It is generally believed that MAC proteins, C6 through C9 arose by a

*Address correspondence and reprint requests to Sylvia L. Smith, Department of Biological Sciences, Florida International University, Miami, Florida 33199, U.S.A. smiths@fiu.edu, Telephone: 1-305-348-3183 (or 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.

series of gene duplications of an ancestral perforin-like gene [8,9], and therefore considered members of the gene family that includes the perforins [10]. MAC proteins share several common structural motifs, such as thrombospondin (TS) type I, low-density lipoprotein receptor class A (LDLR), epidermal growth factor precursor (EGFP) modules and a MAC-perforin domain (MACPF). These modules/domains are conserved in their amphibian and teleost counterparts. This report is the first documenting that they are present in elasmobranch complement proteins. Mammalian C8 is a trimeric oligomer composed of non-identical subunits, alpha (α), beta (β), and gamma (γ) chains, each encoded by a separate gene [11,12]. The gamma subunit belongs to the lipocalin family, and is unrelated to the MAC complement protein family [13] and not essential for MAC lytic activity [14].

The insertion of MAC into the lipid bilayer of target membranes depends on the terminal components (C6 through C9) to undergo considerable conformational change involving hydrophilic-amphiphilic molecular transition with exposure of hydrophobic domains [15]. The binding of C8 to the C5b67 complex that is assembled on the target membrane is via a binding site in the C8 β subunit which is non-covalently bound to the disulfide-linked C8 α and C8 γ subunits. The C8 α chain has a crucial role in completing the assembly of MAC since it rapidly binds and initiates the polymerization of C9 molecules which insert into the lipid bilayer. Partially purified, functionally pure components analogous to human C8 and C9 have previously been described (referred to as C8n and C9n or t1 and t2, respectively) [16,17] indicating that these two significant members of the complex had evolved at the time sharks emerged.

Individual components analogous to mammalian C6 through C9 have only been described for a few non-mammalian species [18–26]. Complement-associated lytic activity has not been described in organisms more basal than the shark [27–30]. A C6-like gene has been cloned from amphioxus, a cephalocordate [31]; whether the encoded protein functions as a component of a MAC in amphioxus remains to be determined.

The nurse shark, *Ginglymostoma cirratum*, is a primitive member of the vertebrate phyla and by virtue of its phylogenetic position serves as an excellent animal model to study ancestral complement genes and proteins since it has elements of innate and adaptive immunity and a complement system with a terminal lytic pathway. Hemolytic activity of shark serum has been known for decades [16,32] and trans-membrane pore structures formed on target membranes have been shown to be structurally similar to those formed by human MAC [3,16], however, the molecular composition of shark MAC has not been determined nor have genes and/or proteins of MAC been cloned or fully characterized. This report is the first to describe the cloning, sequence analysis and expression of the gene encoding the alpha subunit of the shark C8 homologue, the first MAC gene to be cloned from an elasmobranch.

2. Materials and methods

2.1 Materials

TRIzol Reagent, Restriction enzymes, TOPO Cloning Kit, PCR Supermix High Fidelity, Oligo(dT)12–18 primer, and Superscript II Reverse Transcriptase were purchased from Invitrogen Life Technologies (Carlsbad, CA, USA). Big Dye Terminator Cycle Sequencing Kit V2.0 was purchased from Applied Biosystems (Foster City, CA, USA). All gene specific DNA primers were purchased from Sigma-Genosys (St. Louis, MO). The QIAquick Gel Extraction kit was purchased from Qiagen (Valencia, CA, USA). SMARTTRACE cDNA amplification kit was obtained from Clontech (Palo Alto, CA, USA). The Wizard *Plus* SV Minipreps DNA purification system was purchased from Promega (Madison, WI, USA). PCR DIG Probe Synthesis Kit was obtained from Boehringer Mannheim (Indianapolis, IN,

USA), and Lumiphos Plus was purchased from Whatman Biosciences (MA, USA). DNA sequencing reactions were carried out using the Big Dye Terminator V3.1 and the automated sequencer ABI377 (Applied Biosystems, Foster City, CA, USA). TE buffer was composed of 10 mM Tris and 1 mM EDTA and brought to pH 7.5 with HCl. Queen's lysis buffer was composed of 0.01 M Tris, 0.01 M NaCl, 0.01 M disodium-EDTA, and 1.0% n-lauroylsarcosine and brought to pH 8.0.

2.2 Animals

A 2 Kg young female nurse shark (*G. cirratum*) was captured from the waters near the Keys Marine Laboratory, (Long Key, Florida Keys, FL); it was transported to Florida International University for sacrifice and subsequent tissue harvesting. The organism was anesthetized with one part per million of 3-aminobenzoic acid ethyl ester (methane sulfonate) and allowed to bleed out from the caudal vein. After careful dissection the tissues were flash frozen in liquid nitrogen and then stored at -80°C until used in nucleic acid extractions. Whole blood was obtained from adult animals housed in an open sea water channel at the Keys Marine Laboratory in Long Key, Florida.

2.3. First-strand cDNA preparation and degenerate PCR

Total RNA was extracted from homogenized nurse shark liver using the TRIzol Reagent (Invitrogen Life technologies) according to manufacturer's instructions. Using Superscript II reverse transcriptase and Oligo(dT)12–18 primer (Invitrogen Life technologies), first-stranded cDNA was synthesized using 4 μg of total RNA as the template. Degenerate primer NSC9-DGF1 (5' GYAA YGGNGAYAA YGAYTGYG 3') was designed using multiple alignments and based on conserved regions (CNGDQDC, human amino acid positions 115–121 in human C8 α) of human C6, C7, C8 α , C8 β and C9 deduced amino acid sequences and employed in RT-PCR paired with AUAP (Clontech). The amplification program was 94°C for 1 m and 35 cycles of 94°C for 30 s, 56°C for 30 s, and 71°C for 3 m and a final extension of 71°C for 6 m. The resulting DNA fragments of 2.1 and 2.9 kb were detected by electrophoresis and further amplified by nested PCR (SMARTRACE cDNA amplification kit) under thermocycler settings of 94°C for 30 s, 62°C for 30 s, and 72°C for 3 m and 30 s for 25 cycles. The PCR products were run on a 1% agarose gel and then purified using the QIAquick gel extraction kit (Qiagen) according to the manufacturer's instructions. The extracted products were cloned into a TOPO-TA vector (Invitrogen). Recombinants were identified by blue/white colony selection on ampicillin-containing LB agar plates.

2.4 Preparation of Plasmid DNA and Sequencing

Clones positive for the 2.1 Kb insert were selected by colony PCR. Colony PCR amplification was carried out for 30 cycles of 94°C for 30 s, 51°C for 30 s, and 72°C for 2 m. Plasmids of selected clones were purified using the Wizard *Plus* SV Minipreps DNA purification system. The purified plasmids were subjected to cycle sequencing reactions composed of 2 μl Big Dye Terminator V2.1, 2 μl purified plasmid DNA, and 2 μl 0.8 μM primer. Amplification for all sequencing reactions was carried out as follows: initial denaturation at 96°C for 1 m, followed by 28 cycles of 96°C for 5 s, annealing at 50°C for 10 s and final extension at 60°C for 4 m. The resulting PCR products were submitted for sequencing by the automated sequencer ABI377 (Applied Biosystems). C8 α -like clones were sequenced in forward and reverse directions for sequence confirmation. Clones were subjected to cycle sequencing using M13 forward and reverse primers then gene specific primers were constructed from resulting sequences to further sequence the entire gene. All clones overlapped in sequence by at least 100 base pairs. Gene specific primers used to amplify GcC8 α gene are listed in Table 1.

2.5 Amplification and cloning of full-length GcC8 α cDNA

A full length GcC8 α transcript was obtained by long-PCR using primers C8 α -ESFP2 and C8 α -ESRP1 that were designed to the 5' and 3' UTR of the assembled sequence generated from overlapping clones. Amplification was carried out for 38 cycles of: denaturation at 94°C for 30 s, annealing at 60°C for 30 s, and extension at 70°C for 3 m. The PCR mixture was composed of 1 μ l (10 μ M) of each primer, 45 μ l PCR Supermix High Fidelity (Invitrogen Life technologies), and 3 μ l shark liver cDNA. The PCR product was run on a 1% agarose gel with ethidium bromide. The band of expected size was cut out, purified, cloned, and sequenced as described above in section 2.4. Clones positive for the 2.3 Kb insert were selected by colony PCR.

2.6 Sequence and Phylogenetic analysis

The full-length GcC8 α nucleotide sequence was translated to the corresponding amino acid sequence in the BioEdit biological sequence alignment editor for Windows 95/98/NT/2000/XP [33]. The identities of positive clones were established using the Basic Local Alignment Search Tool (BLAST) search engine [34]. Amino acid Identity and similarity percentages were calculated using alignments constructed in ClustalW [35]. Calculations were made by manual counting of identical and similar amino acid residues. Multiple alignments for phylogenetic analysis were constructed by the ClustalX program [36]. This alignment was then used by the PAUP* program [37] to construct a phylogeny using the neighbor joining algorithm [38] under the default settings. Confidence in the branch points were validated by 1000 bootstrap replications. Sequences of other species were obtained from GenBank.

2.7. Molecular analyses

Molecular modules were determined by studying and comparing alignments created by ClustalW of GcC8 α sequence and C8 β sequences of other taxa. Putative N-linked glycosylation sites were predicted by the presence of the amino acid sequence: N(Asp)-X(any amino acid residue)-[S(Ser) or T(Thr)] [39], 1974), where X is followed by a Serine or Threonine residue. Potential mannosylation sites were identified by searching for the sequon W-X-X-W-X-X-W [40]. Hydrophobicity profiles were generated employing the default settings of the Kyte & Doolittle algorithm in the BioEdit program [33].

2.8 Southern blotting

Based on GcC8 α cDNA sequence obtained, the primer set C8ASAZNFP3 and C8ASAZNRP1 was designed to cover a 226 nucleotide sequence for synthesis of a Digoxigenin (DIG) labeled probe using the PCR DIG Probe Synthesis Kit. The template used was first strand cDNA from shark liver. The primers were 19 (sense) and 22 (antisense) nucleotides in length, respectively, and were designed not to extend across introns using the human C8 α intron/exon pattern as a guide [41]. Amplification consisted of 30 cycles of 95°C for 30 s, 54°C for 30 s and 72°C for 30 s. Shark genomic DNA was isolated from whole blood using Queen's lysis buffer in a 1:40 ratio and digested with restriction enzymes *Bam*HI, *Eco*RI, *Hind*III and *Pst*I overnight at 37°C followed by ethanol precipitation. The digested DNA was resuspended in 0.2 X TE buffer and electrophoresed in 0.8% agarose gel at 27 V for nine hours. The DNA was then transferred to a nitrocellulose membrane and fixed by UV cross linking. The membrane was then hybridized with the DIG labeled probe for 16 hr at 42°C. LumiPhos reagent was added to the membrane for chemiluminescent detection of DIG probe and incubated 45 m at 37°C. The blot was exposed for three hours in a dark room onto X-ray film.

2.9. Gene expression

GcC8 α gene expression was determined by RT-PCR. Total RNA was extracted from homogenized nurse shark liver, kidney, brain, intestine, ovary, muscle, heart, pancreas, spleen, erythrocytes and leukocytes using the TRIzol Reagent according to manufacturer's instructions. Using Superscript II reverse transcriptase and Oligo(dT)12–18 primer, first-stranded cDNA was synthesized from each tissue. The resulting nucleic acid was used as a template for RT-PCR. Employing the primers NSC8A-L2 (forward) and C8A-SQRP1 (reverse) (table 1) that span a 503 nucleotide region, RT-PCR was performed under thermocycler settings of 42 cycles of 94°C for 30 s, 54°C for 30 s, and 72°C for 50 s. Universal β -actin specific primers for β -actin (forward: 5'-CTGCCATGTATGTTGCCATC-3' (nucleotide numbers, 389–408) and reverse 5'-ATCCACATCTGCTGGAAGGT-3' (nucleotide numbers 1051–1070) were run simultaneously as a control at the same thermocycler settings except that amplification was carried out for 32 cycles. PCR products were electrophoretically analyzed on a 1% agarose gel containing ethidium bromide.

3. Results

3.1. Cloning and sequence analysis of the full length GcC8 α cDNA

A 2341 nucleotide sequence was constructed from overlapping clones that included the 3' and 5' UTRs. Table 1 lists the primers used. From these overlapping clones a single cDNA sequence was determined. Primers were designed based on the compiled sequence and used in PCR amplification [42] to generate a mRNA transcript representing the full length shark C8 α gene. Several clones representing a single cDNA sequence with high homology to C8 α of the human, mouse, rat, rabbit, and pig were identified. The nucleotides (1770) of the coding region translate into 589 amino acid residues. The full length cDNA sequence and its full translation are displayed in figure 1.

3.2 Multiple Alignment and Sequence Analyses

Using the computer software, ClustalW, the shark C8 α deduced amino acid sequence was aligned with other C8 α amino acid sequences of human, mouse, frog and trout (Figure 2). The shark C8 α sequence contains 33 cysteine residues of those 29 are conserved between the shark and human sequences. These residues are potentially capable of forming, through disulfide bonds, the characteristic C8 α cysteine backbone suggesting a similar folding pattern and function to mammalian, amphibian, and teleost C8 α . The extra three cysteine residues that are not conserved in the human ortholog are at the very beginning of the GcC8 α sequence and are probably part of the leader peptide which does not contribute to C8 α 's membrane attack function. In shark the cysteine residues corresponding to cysteine residue at 164 in human C8 α forms the disulfide bond with Cys⁴⁰ in C8 γ [43], and the cysteines that correspond to C³²⁴ and C³⁴⁹ in human that are proposed to form a disulfide bridge [44] are conserved in the nurse shark C8 α sequence (Figure 2). The multiple alignment also shows that the indel (human aa 159–175) exclusive to C8 α (shark aa 198–207) is present. The indel sequence reported for the trout, based on alignment of all trout MAC aa sequences, shows the trout insertion contains two extra amino acids [20].

Percent amino acid identity and similarity between the nurse shark sequence and C8 α orthologs from other species was calculated from individual alignments of known C8 α sequences and GcC8 α . The average amino acid sequence identity between the nurse shark C8 α gene and other known C8 α sequences was 41.6% identity and 74.1% similarity with the highest identity to its amphibian ortholog at 47.4%.

Analysis of the primary structure showed that GcC8 α has a modular structure consistent with that of C8 α of other taxa examined (Figure 3). The conserved modules identified in GcC8 α are similar to those of mammalian species. The two thrombospondin type 1 repeats, the low density lipoprotein receptor class A repeats, and the membrane attack complex protein/perforin-like segments are present in GcC8 α and are highly conserved.

Four putative N-linked glycosylation sites were identified in the sequence at ASN²⁶, ASN²⁸¹, ASN⁵⁴⁶ and ASN⁵⁵⁸ (Figure 3). The sequence was also examined for potential mannosylation sites, two were identified. One in the first Thrombospondin Type 1 repeat and was in the pattern of: WxxW amino acids 53–56 WAQW. The second site was in the WxxWxxW motif and located toward the end of the sequence in the final Thrombospondin Type 1 repeat domain at amino acids 549–553 in the pattern of WSCWSGW [40]. Kyte and Dolittle [41,45] hydrophobicity analysis comparing GcC8 α to human C8 α were made in the BioEdit program [33] (results not shown). And the profile showed considerable similarity in hydrophobicity between human C8 α and GcC8 α , however, there are differences in distribution of hydrophobic residues particularly within the MAC-PF domain (residues 320-415).

3.3. Phylogenetic analysis

To determine the evolutionary status of GcC8 α a multiple alignment of all known MAC sequences was created in the ClustalX program and without further manipulation of gaps. Entire protein sequences were used to generate the tree. Human, mouse, cow, cat and woodchuck perforin sequences were used as an out group. This alignment was applied in the PAUP* program and a phylogenetic tree constructed (Figure 4) showing that GcC8 α forms a clade with C8 α sequences from other taxa and is sister taxa with frog C8 α . The tree also shows that the C8 complex has diverged from a common ancestor with C9.

3.4. Southern blot and tissue expression analysis of GcC8 α

The expression of GcC8 α gene in tissues of the nurse shark was detected by RT-PCR. This semi-quantitative analysis revealed relatively high levels of C8 α transcripts in the liver which is the primary tissue of complement protein synthesis in most organisms (Figure 5). Surprisingly, C8 α synthesis, albeit at lower levels, was detected in all tissues examined (kidney, brain, intestine, ovary, muscle, heart, pancreas, erythrocytes, and leukocytes), except for the spleen where no expression was detected. β -actin expression, included as a control, was relatively uniform in all tissues examined.

Southern Blot analysis was performed to determine the gene copy number of GcC8 α in the shark genome. A single hybridizing band was detected in each of the shark genomic DNA digests (enzyme digestion by *Bam*HI, *Eco*RI, *Hind*III, and *Pst*I) suggesting that there is a single gene copy of C8 α in the shark genome (Figure 6).

4. Discussion

Lysis of target cells is an effector function of mammalian complement and is accomplished by the assembly and insertion of the membrane attack complex (MAC) into target membranes. To date, no homologue to any one of human MAC genes, with the exception of C5 [46] has been cloned from an elasmobranch [47]. In this study, shark GcC8 α gene was cloned and sequenced. GcC8 α is 2341 nucleotides in length encoding 589 amino acids. Based on the size of the coding region and not taking into account potential glycosylation of the molecule, the predicted molecular weight is likely to be higher than that of human C8 α (554 amino acid residues) which also has fewer N-linked glycosylation sites. An earlier study [16] estimated, from partially purified shark C8, a molecular weight closer to 185kDa, human C8 mature protein is 152kDa. The deduced amino acid sequence of GcC8 α shows

47.4% sequence identity with *Xenopus* C8 α and 41.4% identity with human C8 α . Structural analysis reveals conservation of modules characteristic of mammalian MAC proteins and organized in the same sequential order. The perforin-like segment contains the conserved indel characteristic of C8 α . This conservation of structural similarity further suggests that GcC8 α might be linked non-covalently to a C8 β subunit. It should be noted, however, that homologues of C8 β and C8 γ have yet to be described in the shark. Comparison of the hydrophobicity profile of GcC8 α with that of human C8 α shows consistent similarities in the hydrophobic regions with the exception of regions: 140–190 a region that lies between the LDLR and MACPF modules, 380–395 and 462–470 in the MACPF domain which may influence its insertion into target membrane [48]. The distribution and position of hydrophobic residues through the entire coding region reveals that GcC8 α has the physico-chemical properties to function in a manner similar to C8 α , that is, it most likely participates in hydrophilic-amphiphilic transition and contributes to the assembly and anchoring of a MAC-like macromolecule into target membranes

Four putative N-linked glycosylation sites were identified at positions different from that of human C8 α . Human C8 α has two N-linked glycosylation sites, at ASN⁴³ and ASN⁴³⁹ and only one is suspected to actually be glycosylated [41]. The first N-linked glycosylation site in human C8 α is located in the TSP1 module. A corresponding site in the shark is absent, however, a N-linked glycosylation site is found upstream of the TSP1 module in the leader peptide sequence. The functional significance of this potential site is unclear. The second N-linked glycosylation site is present in the MACPF domain in both human and shark, a region in which there is high sequence conservation between mammal and shark. The remaining two N-linked glycosylation sites in GcC8 α are present in the second TSP1 located at the C-terminal end. Two potential C-mannosylation sites were identified that are highly conserved in all orthologs examined (Fig. 2). The C-mannosylation patterns were: WAQW (aa 53–56) located in the first TSP1 module and WSCWSGW (aa 547–553) in the second TSP1 module at the C-terminal end of the molecule. The location of glycans in the sequence is important since glycosylation can contribute to protein folding and signal response. Glycan structures can interfere with activation site exposure [49]

C8 α is a unique member of MAC in that it contains an indel site that contains the cysteine residue that covalently binds C8 γ [50]. In humans, the indel region is the main sequence that C8 α associates with, even when Cys¹⁶⁴ is replaced by Gly¹⁶⁴ representing sufficient attraction to bind non-covalently [50]. Multiple alignment (Fig. 2) shows that the corresponding cysteine as well as the region corresponding to human indel is conserved indicating that GcC8 α is a C8 α ortholog. The shark indel is located between the LDLR and the MACPF domain and contains the conserved cysteine residue at position 203 that in human C8 α forms the disulfide bond with C8 γ suggesting that GcC8 α occurs as a disulfide linked α - γ dimer. This region is highly conserved between human and shark C8 α (Fig. 3) showing 88.2 % identity and 94.1% similarity to the human indel sequence. Whether the α -subunit non-covalently links to a β subunit forming a trimeric molecule such as the C8 of mammals and teleosts, remains to be confirmed.

All cysteine residues of human C8 α form intra-molecular disulfide bonds with the exception of Cys¹⁶⁴ that forms an inter-molecular disulphide bond with Cys⁴⁰ in C8 γ [51]. The MAC proteins are rich in cysteine residues and the multiple alignment (Fig. 2) demonstrates that a total of 29 cysteines are conserved between elasmobranch and mammal. GcC8 α , however, is more cysteine rich as it has four extra cysteine residues located toward the N-terminal of the sequence in the leader peptide region and are not likely to be involved in GcC8 α function. Human MAC proteins and perforin have two important conserved cysteine residues (C7: Cys³¹⁷ and Cys³³³, C9: Cys³⁵⁸ and Cys³⁸³, C8 α : Cys³⁴⁶ and Cys³⁷⁰, C8 β : Cys³²⁴ and Cys³⁴⁹, perforin: Cys²³⁶ and Cys²⁵⁸) that form a disulfide bridge. The loop

formed by this bond is suspected to be outside the membrane when the transmembrane molecule (i.e., human MAC) is inserted [44]. These two pertinent cysteine residues are present in GcC8 α (highlighted in black in figure 2) and suggest similar functional role.

In several teleost species some complement genes are present in several isoforms [24, 56–61]. Similarly, in elasmobranchs certain complement genes are present in multiple forms, such as, GcC3-1 and GcC3-2, (Smith, unpublished), GcBf/C2-1 and -2 [62] and GcIf-1, -2, -3 and -4 [63] in the nurse shark and TrscBf-A and -B in the banded houndshark (*Triakis scyllia*). Southern blot analysis with a probe corresponding to a region that did not overlap with C8 β (should such a homologue be present in the shark) showed that there is a single gene copy of GcC8 α . Gene expression studies revealed GcC8 α is synthesized in several tissues including erythrocytes and leukocytes, with highest expression in the liver. Interestingly, the expression in peripheral blood cells is higher in erythrocytes than in leukocytes. This indicates that nucleated erythrocytes of shark are transcriptionally active. Multiple organ/cell expression of C8 α is not seen in mammals, where C8 α is primarily synthesized in the liver, however, in other vertebrate species such as trout C8 α and C8 β are expressed in several tissues [18,21]. Taken together these observations suggest that poikilothermic vertebrates synthesize complement proteins more ubiquitously than mammals. As the complement system evolved the tissue sites for complement synthesis may have reduced through evolution, however, the liver (hepatopancreas in some species) remains likely the main site of complement protein synthesis in vertebrates.

Phylogenetic analysis of GcC8 α sheds insight on the evolution of the MAC family of proteins. There are two main theories that attempt to explain the evolution of this significant gene family. Phylogenetic analyses by Mondragon-Palomino et al [9] using MAC amino acid sequences supports the view that C6 and C7 are of earlier origin followed by the emergence of C8 then C9. This group does not present data as to the whether C8 α or β are of earlier derivation. They also prescribe that the terminal complex proteins (C6 through C9) originated from a single ancestral gene composed of complex modular structure. They state that a series of gene duplications and loss of structural modules resulted in complement proteins that make up the MAC protein family. In contrast an earlier hypothesis proposed by Podack et al [64] conjectures that due to the similarity of function, size and sequence, C9 emerged first from a gene duplication event of an ancestral gene common to both perforin and the MAC family proteins. This group further speculates that following C9 emergence C8, C7 and C6 successively emerged through later gene duplication events and developed increasing modular complexity and size. A more recent somewhat different hypothesis that also supports the C8/C9 faction as originators of MAC is suggested by Kauffman et al, [65]. They state that after distance analyses of human MAC components and perforin C8 α and β have a closer phylogenetic distance to perforin than to C6, C7, and C9, maintaining that MAC arose from a fundamental C8-like building block. The phylogenetic analysis performed in this study supports the hypothesis that C8 and C9 are derived from a common ancestor and represent an early duplication event that most likely predated C6 and C7 [56]. Although C6-like molecules have been described for *Amphioxus* [31] and *Ciona* [65], their role as complement proteins remains unconfirmed. Molecular analysis of the C6-like gene described for *Amphioxus* reveals a 5' C6 modular structure with a 3' end missing key modules characteristic of C6. This could also be interpreted to be an early C8-like molecule before loss of the extra TSP1 module at the 5' end. Furthermore, in *Ciona*, the C6-like gene is expressed as a cell-surface receptor and it is unknown whether it has complement function [65,66]. Complement-related lysis has not been detected in either organism. C8 and C9 are the only MAC components (with the exception of C5) that have been detected in the shark. Functional and molecular studies and Western blot analysis of shark MAC proteins have provided no evidence for C6 and C7 to date. The shark, being the most basal organism with

a fully functioning MAC suggests that these proteins are particularly significant members of the membrane attack complex and potentially the first to have evolved.

Acknowledgments

We thank Penelope, the shark, for donating her organs for the project, the FIU DNA Sequencing Core for their services and the Keys Marine Laboratory for housing, maintaining and bleeding the sharks. LA was supported by the MBRS RISE program (R25 GM061347). The work was supported by a student summer research award to LA from the Biomedical Research Initiative (GM061347) and by grant to SLS (GM08205). We also thank the FIU Comparative Immunology Institute for providing research facilities. This report is contribution FIU-CII-006 from the FIU CI Institute. The sequence described in this paper has been deposited in the GenBank data base under accession number EF654112. The work was conducted with institutional IACUC approval.

Abbreviations and definitions

aa	amino acid
AUAP	abridged universal amplification primer
BLAST	Basic Local Alignment Search Tool
C	complement
C1	C2, ...C9, complement components 1 through 9
EGFP	epidermal growth factor precursor
GcC8α	<i>Ginglymostoma cirratum</i> complement component alpha
GcIf	<i>Ginglymostoma cirratum</i> factor I
IACUC	Institutional Animal Care and Use Committee
indel	insertion and/or deletion site
LB	Luria Bertani
LDLR	low-density lipoprotein receptor class A
m	minute
MAC	membrane attack complex
MACPF	MAC-perforin segment
PAUP	Phylogenetic Analysis Using Parsimony
RACE	rapid amplification of cDNA ends
RT-PCR	Reverse Transcriptase-PCR
s	second
TE	Tris-EDTA
TS	thrombospondin type I
UTR	untranslated region

References

1. Walport MJ. Advances in Immunology: Complement part 1. New England Journal of Medicine. 2001a; 344:1058–1066. [PubMed: 11287977]
2. Walport MJ. Advances in Immunology: Complement part 2. New England Journal of Medicine. 2001b; 344:1140–1144. [PubMed: 11297706]

3. Humphrey JH, Dourmashkin RR. The lesions in cell membranes caused by complement. *Advances in Immunology*. 1969; 11:75–115. [PubMed: 4908393]
4. Muller-Eberhard HJ. The membrane attack complex of complement. *Annual Review of Immunology*. 1986; 4:503–528.
5. Esser AF. The membrane attack complex of complement: Assembly, structure and cytotoxic activity. *Toxicology*. 1994; 87:229–247. [PubMed: 8160186]
6. Thompson RA, Lachmann PJ. Reactive lysis: the complement-mediated lysis of unsensitized cells. I. The characterization of the indicator factor and its identification as C7. *Journal of Experimental Medicine*. 1970; 131(4):629–41. [PubMed: 4193934]
7. Podack ER, Tschopp J. Membrane attack by complement. *Molecular Immunology*. 1984; 7:589–603. [PubMed: 6379417]
8. Young JD, Liu CC, Leong LG, Cohn ZA. The pore-forming protein (perforin) of cytolytic T lymphocytes is immunologically related to the components of membrane attack complex of complement through cysteine-rich domains. *Journal of Experimental Medicine*. 1986; 164(6):2077–82.
9. Mondragon-Palomino M, Pinero D, Nicholson-Weller A, Lactette JP. Phylogenetic Analysis of the Homologous Proteins of the Terminal Complement Complex Supports the Emergence of C6 and C7 Followed by C8 and C9. *Journal of Molecular Evolution*. 1999; 49(2):282–9.
10. Hobart MJ, Fernie BA, DiScipio RG. Structure of the human C7 gene and comparison with the C6, C8A, C8B, and C9 genes. *Journal of Immunology*. 1995; 154(10):5188–94.
11. Steckel EW, York RG, Monahan JB, Sodetz JM. The eighth component of human complement: purification and physico-chemical characterization of its unusual subunit structure. *Journal of Biological Chemistry*. 1980; 255:11997–12005. [PubMed: 7440581]
12. Ng SC, Rao AG, Howard OM, Sodetz JM. The eighth component of human complement (C8): evidence that it is an oligomeric serum protein assembled from products of three different genes. *Biochemistry*. 1987; 26:5229–5233. [PubMed: 3676249]
13. Haefliger JA, Peitsch MC, Jenne DE, Tschopp J. Structural and functional characterization of complement C8 gamma, a member of the lipocalin protein family. *Molecular Immunology*. 1991; 28:123–31. [PubMed: 1707134]
14. Brickner A, Sodetz JM. Function of subunits within the eighth component of human complement: selective removal of the γ chain reveals it has no direct role in cytolysis. *Biochemistry*. 1984; 23:832–837. [PubMed: 6712927]
15. Sodetz JM. Structure and Function of C8 in the membrane attack sequence of complement. *Current Topics in Microbiology and Immunology*. 1989; 140:19–31. [PubMed: 2644073]
16. Jensen JA, Festa E, Smith DS, Cayer M. The complement system of the nurse shark: hemolytic and comparative characteristics. *Science*. 1981; 214:566–569. [PubMed: 7291995]
17. Smith SH, Jensen JA. The second component (C2n) of the nurse shark complement system: purification, physico-chemical characterization and functional comparison with guinea pig C4. *Developmental and Comparative Immunology*. 1986; 10:191–206.
18. Kazantzi A, Sfyroera G, Holland MG-H, Lambris JD, Zarkadis IK. Molecular cloning of the beta subunit of complement component eight of the rainbow trout. *Developmental and Comparative Immunology*. 2003; 27:167–174. [PubMed: 12590968]
19. Kalagiri T, Hirono I, Aoki T. Molecular analysis of complement component of C8 β and C9 cDNAs of Japanese founder, *Paralichthys olivaceous*. *Immunogenetics*. 1999; 50:43–48. [PubMed: 10541805]
20. Papanastasiou AD, Zarkadis IK. The gamma subunit of the eighth complement component (C8) in rainbow trout. *Developmental and Comparative Immunology*. 2006; 30(5):485–91. [PubMed: 16169077]
21. Papanastasiou AD, Zarkadis IK. Cloning and phylogenetic analysis of the alpha subunit of the eighth complement component (C8) in rainbow trout. *Molecular Immunology*. 2006; 43(14):2188–94. [PubMed: 16483657]
22. Uemura T, Yano T, Shiraishi H, Nakao M. Purification and characterization of the eighth and ninth components of carp complement. *Molecular Immunology*. 1996; 33:925–32. [PubMed: 8960116]

23. Chondrou MP, Mastellos D, Zarkadis IK. cDNA cloning and phylogenetic analysis of the sixth complement component in rainbow trout. *Mol Immunol*. 2006; 43:1080–1087. [PubMed: 16271768]
24. 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; 54:807–815. [PubMed: 12618914]
25. Zarkadis IK, Duraj S, Chondrou M. Molecular cloning of the seventh component of complement in rainbow trout. *Developmental and Comparative Immunology*. 2005; 29:95–102. [PubMed: 15450749]
26. Papanastasiou AD, Zarkadis IK. The gamma subunit of the eighth complement component (C8) in rainbow trout. *Developmental and Comparative Immunology*. 2006; 30:485–491. [PubMed: 16169077]
27. Fujii T, Murakawa S. Immunity in lamprey. III. Occurrence of the complement-like activity. *Developmental and Comparative Immunology*. 1981; 5(2):251–9. [PubMed: 6786926]
28. Smith LC, Azumi K, Nonaka M. Complement system in invertebrates. The ancient alternative and lectin pathways. *Immunopharmacology*. 1999; 42:107–120. [PubMed: 10408372]
29. Dishaw LJ, Smith SL, Bigger CH. Characterization of a C3-like cDNA in a coral: phylogenetic implications. *Immunogenetics*. 2005; 57(7):535–48. [PubMed: 16041540]
30. Nonaka M, Smith SL. Complement system of bony and cartilaginous fish. *Fish & Shellfish Immunology*. 2000; 10:215–28. Review. [PubMed: 10938735]
31. Suzuki NM, Satoh N, Nonaka M. C6-like and C3-like molecules from the cephalochordate, amphioxus, suggest a cytolytic complement system in invertebrates. *Journal of Molecular Evolution*. 2000; 54:671–679. [PubMed: 11965439]
32. 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]
33. 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.
34. Altschul S, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of Molecular Biology*. 1990; 215:403–410. [PubMed: 2231712]
35. 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; 22:4673–80. [PubMed: 7984417]
36. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*. 1997; 25:4876–82. [PubMed: 9396791]
37. Swofford, DL. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer Associates; Sunderland, Massachusetts: 2002. PAUP*. Version 4
38. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*. 1987; 4:406–25. [PubMed: 3447015]
39. Marshall RD. The nature and metabolism of the carbohydrate-peptide linkages of glycoproteins. *Biochemical Society Symposia*. 1974; 40:17–26.
40. Hofsteenge J, Blommers M, Hess D, Furmanek A, Miroshnichenko O. The four terminal components of the complement system are C-mannosylated on multiple tryptophan residues. *Journal of Biological Chemistry*. 1999; 274:32786–94. [PubMed: 10551839]
41. Morley, BJ.; Walport, MJ. *The Complement Facts Book*. Academic Press; 2000. p. 123-130.
42. Frohman MA, Dush MK, Martin GR. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Sciences*. 1988; 85(23):8998–9002.
43. Brickner A, Sodetz JM. Functional domains of the α subunit of the eighth component of human complement: identification and characterization of a distinct binding site for the γ chain. *Biochemistry*. 1985; 24:4603–7. [PubMed: 4063342]

44. Peitsch MC, Amiguet P, Guy R, Brunner J, Maizel JV, Tschopp J. Localization and molecular modelling of the membrane-inserted domain of the ninth component of human complement and perforin. *Molecular Immunology*. 1990; 27:589–602. [PubMed: 2395434]
45. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. 1982; 157:105–32.
46. Graham M, Shin D-H, Smith SL. Molecular and expression analysis of complement component C5 in the nurse shark (*Ginglymostoma cirratum*) and its predicted functional role. *Fish and Shellfish Immunology*. 2009;10.1016/j.fsi.2009.04.00
47. Smith SL. Shark complement: an assessment. *Immunological Reviews*. 1998; 166:67–78. [PubMed: 9914903]
48. Hadders MA, Berenger DX, Gros P. Structure of C8 α -MACPF reveals mechanism of membrane attack in complement immune defense. *Science*. 2007; 317:1552–1554. [PubMed: 17872444]
49. Wells L, Vosseller K, Hart GW. Glycosylation Nucleocytoplasmic Proteins: Signal Transduction and O-GlcNAc. *Science*. 2001; 291:2376–8. [PubMed: 11269319]
50. Plumb ME, Sodetz JM. An indel within the C8 α subunit of human complement C8 mediates intracellular binding of C8 γ and formation of C8 α - γ . *Biochemistry*. 2000; 39:13078–13083. [PubMed: 11041874]
51. Slade DJ, Chiswell B, Sodetz JM. Functional studies of the MACPF domain of human complement protein C8 α reveal sites for simultaneous binding of C8 β , C8 γ , and C9. *Biochemistry*. 2006; 45:5290–6. [PubMed: 16618117]
52. Sunyer JO, Zarkadis IK, Sahu A, Lambris JD. Multiple forms of complement C3 in trout that differ in binding to complement activators. *Proceedings of the National Academy of Sciences*. 1996; 93:8546–51.
53. Sunyer JO, Tort L, Lambris JD. Structural C3 diversity in fish: characterization of five forms of C3 in the diploid fish *Sparus aurata*. *Journal of Immunology*. 1997; 158:2813–21.
54. Nakao M, Fushitani Y, Fujiki K, Nonaka M, Yano T. Two diverged complement factor B/C2-like cDNA sequences from a teleost, the common carp (*Cyprinus carpio*). *Journal of Immunology*. 1998; 161:4811–8.
55. Kuroda N, Wada H, Naruse K, Simada A, Shima A, Sasaki M, Nonaka M. Molecular cloning and linkage analysis of the Japanese medaka fish complement Bf/C2 gene. *Immunogenetics*. 1996; 44:459–67. [PubMed: 8824158]
56. Papanastasiou AD, Zakardis IK. Gene duplication of the seventh component of complement in rainbow trout. *Immunogenetics*. 2005; 57:703–708. [PubMed: 16142474]
57. Nakao M, Matsumoto M, Nakazawa M, Fujiki K, Yano T. Diversity of complement factor B/C2 in the common carp (*Cyprinus carpio*): three isotypes of B/C2-A expressed in different tissues. *Dev Comp Immunol*. 2002; 26:533–541. [PubMed: 12031413]
58. Sunyer JO, Zakardis I, Sarrias MR, Hansen JD, Lambris JD. Cloning, structure and function of two rainbow trout Bf molecules. *Journal of Immunology*. 1998; 161:4106–14.
59. Chondrou M, Papanastasiou AD, Spyroulias GA, Zakardis IK. Three isoforms of complement properdin factor P in trout? Cloning, expression, gene organization and constrained modeling. *Developmental and Comparative Immunology*. 2008; 32:1454–1466. [PubMed: 18638502]
60. Gongora R, Figueroa F, Klein J. Independent duplications of Bf and C3 complement genes in the zebrafish. *Scandinavian Journal of Immunology*. 1998; 48:651–658.
61. Papanastasiou AD, Georgaka E, Zakardis IK. Cloning of CD59-like gene in rainbow trout. Expression and phylogenetic analysis of two isoforms. *Molecular Immunology*. 2007; 44:1300–1306. [PubMed: 16876248]
62. Shin DH, Webb B, Nakao M, Smith SL. Molecular cloning, structural analysis and expression of complement component Bf/C2 genes in the nurse shark, *Ginglymostoma cirratum*. *Developmental and Comparative Immunology*. 2007; 31:1168–82. [PubMed: 17482263]
63. Shin D-H, Webb M, Nakao M, Smith SL. Characterization of shark factor I gene(s): geneomic analysis of novel shark-specific sequence. *Molecular Immunology*. 2009;10.1016/j.molimm.2009.04.002
64. Podack ER, Hengartner H, Lichtenheld MG. A central role of perforin in cytotoxicity? *Annual Review of Immunology*. 1991; 9:129–147.

65. Kauffman DL, Keller PJ, Bennick A, Blum M. Alignment of amino acid and DNA sequences of human proline-rich proteins. *Critical Reviews in Oral Biology and Medicine*. 1993; 4:287–92. [PubMed: 8373986]
66. Wakoh T, Ikeda M, Uchino R, Azumi K, Nonaka M, Kohara Y, Metoki H, Satou Y, Satou M, Sataka M. Identification of transcripts expressed preferentially in hemocytes of *Ciona intestinalis* that can be used as molecular markers. *DNA Research*. 2004; 11:345–352. [PubMed: 15747582]
67. Azumi K, Santis RD, Tomaso AD, Rigoutsos I, Yoshizaki F, Pinto MR, Marino R, Shida K, Ikeda M, Ikeda M, Arai M, Inoue Y, Shimizu T, Satoh N, Rokhsar DS, Pasquier LD, Kasahara M, Satake M, Nonaka M. Genomic analysis of immunity in a Urochordate and the emergence of the vertebrate immune system: “waiting for Godot”. *Immunogenetics*. 2003; 55:570–581. [PubMed: 14530883]

1 AAAAAATATTATTACACTGCATGAAGAATGAAT**ATG**TTTTGAGCTGCTTCTCAACTTTTGCATGCTGCTGTTTTTACTCATTITGCC 90
M F L S C F S T F A C L L F F T H F A

91 AATTCGAATCAGATCAGCAACAGTTCAAACATTGTCACGCATCAGTCTGGCCGAAAAATCCGAAGTAGTCCCACTTGATGTCAA 180
N S N Q I S N S S K H C C T H Q S G R K I R S V V P L D C Q

181 CTTGGCCAGTGGGCACAATGGACAGTGTGTTCTCCTTTGGAAGCAAGAAAGTCCAGATATAGGAATCTGAATCGGCCAGCATATATGGT 270
L G Q W A Q W T V C S P C E A R K S R Y R N L N R P A I Y G

271 GGTAGCCAATGCATTGGATCTCTCTGGGAAGACACATCTGTGAAACATCTGAACAATGCATACCAAAATTAACGTGGTGATCAGTT 360
G S Q C I G S L W E D T S C E T S E Q C I P K I N C G D Q F

361 CAATGTAGTTCCAGTTCAGTATTAATAAGACATTTACTGTGCAATGGTGAGAAAGACTGTGCAGATGTTTCAGATGAAGAAACCTGTGAG 450
Q C S S G R C I K R H L L C N G E K D C A D V S D E E T C E

451 TCTGATTATCCATATGAGAGAAGGACTTTCTGCAGTGACTTTTTTAATCCAGGCATAGAAGCAATTATGACTGGCTATAACATCTTA 540
S D Y P Y E R R T F C S D L F L I P G I E A I M T G Y N I L

541 ATACATGACGTTGGAAGAACTGTGTAGACACTGGGTTTGGTGGATATTGTGAATATGTGTATTAATGGGACTGGAGGACTGCGATAT 630
I H D V G R T V L D T G F G G Y C E Y V Y N G D W R E L R Y

631 GACAGTGAGTGTGAACGCCCTCTACTATAACGATGATGAGAATACTACCGAAACCATACAATTACTCACATACCGTTTGGAGCAATT 720
D S E C E R L Y Y N D D E K Y Y R K P Y N L L T Y R F E A I

721 GCTGATTCTGGATTACCATAGATGTTCAACATGATGTACATGAATTAATTACAGCAATGAAGCATACAGACTCCTTTGATTGTCAGTC 810
A D S G F T I D V H N D V H E L I T A M K H T D S F D F A V

811 GGTGTGCAAGTTTCTGTGTAAGAACTGCTATTGCTTTGGAGAATATCCACATTTATTAGAAATGTATCAAGTTTCAAGGGAAGGAT 900
G V Q V S V V K A A I A F G E Y S T F I R N V S R F E A I

901 GTCAGTTTGTGCTGTGCTACTAAGATTCAAAGTCTCACTTCAAAATGAGGAGATATAATTACTTTTGGATGAAGATATGACTCAA 990
V S F V R V R T K I Q T A H F K M R R Y N L L L D E D M T Q

991 TCCTCATGAACTTCAGATGAATACAATTACGGAATGTACGCTAAATTCATTCAGATTATGGTACTCATTACTGCTTCAGGAACA 1080
S L M K L P D E Y N Y G M Y A K F I A D Y G T H Y Y A S G T

1081 ATGGGAGGTGTTTATGAGTTCACTTGTCTGTAATAAGAGCATTGGCAGAGTCAGATTTAAGTGCATCCGAAGCTGGATTGCGCTT 1170
M G G V Y E F I L V L N K K A L A E S D L T A S E A G F C V

1171 GCAGGATCACTGGGAGTAGTGGTTTCAAAGGAACCATGGATGTGGTGGCTCGGTAAGGCCAAAGGTTGTAGACGTAATTTGGAGCAG 1260
A G S L G V V V S K G T M D V G G S V K A K G C R R K L E Q

1261 AATGAACCTTCAGAAAGTCAAGAAGCCTTGTGAAGATGTTCTTCCACATGACTGGGAGGTGACCTGAAATCAAGTGTGGACTATTA 1350
N E P S E K S R S L V E D V L P H V L G G D L K S S A G A L L

1351 GGTCAATGGAATACCAAGATGTCAAATGTATCGGCATTGGGGTAAATCTTTGAAATACCTGCCAGCTGCTATTGACTTTGAGCTCATGCCA 1440
G H G I P D V K M Y R H W G K S L K Y L P A V I D F E L M P

1441 ATTTATGACTTGGTGGCAGAGTAACCTTCAGTCTGTTGAATCAACACAAAATTTAAAGAGCCATGGAAGAGTACTTGGTGGAG 1530
I Y D L V A R S K L Q S V E I K Q Q N L K R A M E E Y L V E

1531 TTTTCATCCCTGTAGTGTCTGTGTGCTATAATAATGAAAGCAGTCTTCTGGACAATGTCTGTACCTGTGAATGTCTGCGCTGGTTAT 1620
F H P C R C P V C Y N N G K A V L L D N V C T C E C L P G Y

1621 GAAGGGAAGCGTGCCAGGATACTAAAGAAAGGGGCCCACTAATGGAATTGGAGCTGTTGGTCAGGGTGGTCTTCTGCCAGAATGGT 1710
E G K A C Q D T K R K G P T N G N W S C W S G W S S C Q N G

1711 TCAAAACAGCGAACGAGCTGTAACCATCCACCCCTAAAGATGGTGGTGTACATGCTTGGGAAGAACACACAACCAAGCATGTT 1800
S K Q R T R S C N H P P P K D G G A T C L G K N T Q T K H C

1801 **TAA**ACAAAAATAATACCTCTGGCATTGTCGCGCTTGAATTTTGGCTTTGGGTGATCATTATTTCTTTGAGTATAGATGAATATGAG 1890

1891 GTATAGGCAACATTTTGAATACTACTCTGTGCTTTGTTGATTCTTAAATGTAATGAATGTTGAGATGTATTCAATGATGTGAACA 1980

1981 TTTGAAGTACGCTGACAGAATATCCTTTATTAAGAAGATTAAGCTGATTTGTGTGATTAATGTTTGGCCAGGAAGCTGTGTCAC 2070

2071 ACTTCCAGGACCATCATTACAGTGTGTTGGTCAACACTGCAATGAATATGGGAGTGCAGATTTGAATAGACCTCATTGTGATGAAT 2160

2161 ACATGTTGACCTTGATATCTGATCCATTGTTTCTTTACAGCAGGAACAGATGTTAGTAAGAAAATATGAAAATCAATATATATTT 2250

2251 AGCAGC**TTAAAT**TTCTGTACAGGATATCAAGCAGTTATGATTCCTATTACT**TAATAA**GAGATGCC**AAAAAAAAAAAAAAAAAAAA** 2341

Figure 1.

Nucleotide and deduced amino acid sequence of GcC8a cDNA. The nucleotide sequence is above and the deduced amino acid sequence is below. Underlined bold letters indicate initiation codon, stop codon, and polyadenylation recognition signal, and the polyadenylation tail sequence. Putative N-linked glycosylation sites are indicated by bold, italicized N's and mannosylation sites indicated by bold, italicized sequences beginning and ending with W.

Figure 2.

Full length amino acid sequence alignment in ClustalW of GcC8α with homologues from other organisms: *Homo sapiens*, *Mus musculus*, *Xenopus tropicalis*, and *Oncorhynchus mykiss*. Residues that are identical in all compared proteins are designated by an asterisk. Conservation of strong groups is indicated by semi-colon (;) and periods (.) indicate conservation of weak groups. The conserved cysteine residue that potentially forms the disulfide bond with C8γ is indicated by γ. Potential N-linked glycosylation sites are indicated by bold and double strikethrough. Potential C-mannosylated sites are underlined. Cysteines that are conserved between mammalian C8α and GcC8α are indicated by triangle below. The insertion/deletion region in the sequences is highlighted in bold italics. Two conserved cysteine residues that are predicted to form a disulfide bridge and are conserved throughout all MAC proteins are also highlighted in outlined bold and shadow. The number of corresponding amino acid sequence is given on the right ends.

Figure 3.

Modular and glycosylation site map comparison of human and shark C8α. Schematic representation of the organization of modules/domains 5' – 3' direction in GcC8α shows it to be similar to that of human C8α in domain architecture: thrombospondin type 1 repeats (squares), a low density lipoprotein receptor class A repeat (diamond), membrane attack complex protein/perforin-like segment (long rectangle), and an epidermal growth factor region (oval). Potential N-linked glycosylation sites are featured as stalk-circle with an N and predicted C-mannosylation sites as a stalk square with a C. The GcC8α indel was aligned in ClustalW with the human C8α indel. Asterisk (*) denote residue identity and period (.) indicate similar residues.

Figure 4.

Phylogenetic analysis – Phylogenetic analysis of MAC amino acid sequences across taxa. Alignment completed in ClustalX and phylogenetic tree generated by PAUP rooted on perforin from three taxa. Accession numbers of the sequences used to construct the tree are as follows: Human C6 BAD02321, Trout C6 CAF22026, Frog C6 AAH76972, Amphioxys C6 BAB47147, Human C7 CAA60121, Trout C7-1 CAD92841, Trout C7-2 CAF22025, Mouse C7 XP_356827, Human C8a NP_000553, Chimpanzee C6, NP_001009015, Orangutan C6 BAD02323, Dog C6-1 XP_868028, Dog C6-2 XP_536488, Chicken C6 XP_429140, Xenopus C6 AAH76972, Zebrafish C6 NP_956932, Pig C7 NP_999447, Rat C7 XP_226803, Chicken C7 XP_424774, Flounder C7 BAA88899, Shark C8a EF654112, Rabbit C8a AAA31191, Chicken C8a XP_426667, Xenopus C8a AAH74554, Trout C8a CAH65481, Human C8a NP_000553, Human C8β NP_000057, Cow C8β-1 XP_590870, Cow C8β-2 XP_870144, Dog C8β XP_536694, Mouse C8β NP_598643, Trout C8β AAL16647, Flounder C8β BAA86877, Human C9 NP_001728, Trout C9 P06682, Fugu C9 AAC60288, Rat C9 NP_476487, Mouse C9 NP_038513, Zebrafish C9 NP_001019606, Grass Carp C9 AAS76086, Killifish C9 AAR87007, Flounder C9 BAA86878, Trout C9 CAJ01692, Dog C9 XP_536494, Human perforin AAA60065, Mouse perforin CAA42731, Cow AAQ82904, Cat perforin NP_001095130, and Woodchuck perforin AAG24611.

Figure 5.

Tissue expression of GcC8α. RT-PCR analysis of GcC8α expression in shark kidney (K), spleen (S), brain (B), liver (L), intestine (I), ovary (O), muscle (M), heart (H), pancreas (P), red blood cells (R), and white blood cells (W). The GcC8α gene was expressed at some level in all tissues/cells examined with the exception of the spleen where no expression was detected. The highest gene expression was observed in the liver. Expression of β-actin in tissues was amplified as the control.

Figure 6.

Southern blot analysis. Genomic DNA from shark whole blood was isolated and digested with the restriction enzymes, *Bam*HI (Lane 2), *Eco*RI (Lane 3), *Hind*III (Lane 4), and *Pst*I (Lane 5), electrophoresed, transferred to a nylon membrane and subjected to hybridization with a DIG-labelled probe. A lambda/*Hind*III ladder was run in lane 1, the scale is displayed to the right of the blot.

Table 1

Primers used for sequence analysis, synthesizing PCR-DIG probes and RT-PCR analysis of the GcC8 α gene

Primer Name	Sequence of primer (5'→3')	location in sequence*
NSC9-DGF1	GAAAYGGNGAYAAAYGAYTGYG	115–121
C8ASAZNFP3	CAAAACAGCGAACACGAAGC	1712–1731
C8ASAZNRP1	GAAATCAACAAAGAACACAGAG	1916–1937
NSC8A-L2	GCCGAAAAATCCGAAGTGTA	143–162
NSC8A-I3	GACTGGAGGGAACTGCGATA	610–629
C8A33F	AGGCATTGGCAGAGTCAG	1121–1138
C8AFP10	TGTCTGCCTGGTTATGAAGG	1606–1625
C8A-seqRP1	CTGGACTTTCTTGCTTCAC	218–236
C8A-SQRP1	TGGTTTTCGGTAGTATTTCTC	667–687
C8A-SQRP2	TTACCGAGCCACCCACA	1212–1228
C8A-SQRP3	GGCACGCTTCCCTTCAT	1619–1636
C8a-ESFP2	ATTACACTGCATGAAGAATGA	11–31
C8a-ESRP1	CTGGTAATGATGGACCTGG	2075–2093

* indicates the amino acid position in human C8 α chain