Fish & Shellfish Immunology 40 (2014) 414–423

Contents lists available at ScienceDirect

Fish & Shellfish Immunology

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

Short sequence report

Molecular characterization and expression analysis of the complement factor I (CpFI) in the whitespotted bamboo shark (*Chiloscyllium plagiosum*)

Ying Wang [a], Biao Chen [a], Yan Ke [b], Conghui Wang [a], Boping Ye [a, *]

[a] School of Life Science and Technology, China Pharmaceutical University, Nanjing 210009, PR China
[b] National Center for Traditional Chinese Medicine, Beijing 100027, PR China

ARTICLE INFO

Article history:
Received 15 April 2014
Received in revised form
11 July 2014
Accepted 25 July 2014
Available online 7 August 2014

Keywords:
Complement
Factor I
Regulatory protein
Shark
Chiloscyllium plagiosum

ABSTRACT

Complement factor I (FI) is a plasma serine proteinase that plays an essential role in the modulation of the complement cascade. In the presence of substrate modulating cofactors (Factor H, C4bp, CR1, etc), FI cleaves the activation products of C3 (i.e. C3b) and C4 (i.e. C4b) to limit complement activity. In this study, the full length cDNA of factor I (CpFI) is isolated from the liver of the whitespotted bamboo shark (*Chiloscyllium plagiosum*). The CpFI cDNA is 2326 bp in length, encoding a protein of 671 amino acids, which shares 72–80% identity with FI molecules of other sharks, higher than the teleosts (37–40%) and mammals (44–47%). The sequence alignment and comparative analysis indicates the FI proteins are well conserved, with the typical modular architecture and identical active sites throughout vertebrate evolution, suggesting the conserved function. However, the additional sequence present between the leader peptide (LP) and the factor I membrane attack complex (FIMAC) domain in other fishes is also found in CpFI, which consists of two kind of tandem repeats. Phylogenetic analysis suggests that CpFI belongs to the elasmobranch clade, in parallel with the higher vertebrates, to form a sister taxa to teleosts. Expression analysis revealed that CpFI is ubiquitously distributed in a variety of tissues, with the constitutive expression in liver, which might reflect the species-specific distribution patterns of FI. Together with earlier reports, the presence of FI in various sharks might suggest the existence of a well-developed complement regulation mechanism in cartilaginous fish.

1. Introduction

The complement system is a major recognition and effector mechanism of innate immunity, enabling efficient host-defence against pathogens. It consists of a complex group of serum proteins, glycoproteins and soluble or membrane-bound receptors [1], in which the third component of complement, C3, plays a pivotal role [2]. Specifically, activation of C3 is the central step in the complement activation cascade, which is initiated by three different pathways—the classical, lectin and alternative pathways, thereby orchestrates and connects various immune reactions [3,4]. Actually, key to the prompt responsiveness of complement is spontaneous, constant cleavage of C3 to C3b at low rate [5], also called as “tick-over”, which leads to rapid initiation of a proteolytic complement cascade, as well as potentially presents serious threat to host [6]. However, the presence of complement inhibitors that tightly regulate complement cascades is an essential strategy for host protection against indiscriminate immune surveillance [7,8].

Complement factor I (FI) is a soluble serine protease that regulates multiple pathways in complement activation, by proteolytic cleavage and degradation of C4b and C3b in presence of cofactors [9–11]. Thus, it prevents the assembly of the C3 and C5 convertase enzymes, as well as the downstream effects, avoiding inappropriate amplification of the host immune response [12,13]. In addition, it generates a more effective opsonin (iC3b) from C3b, which is important for the biological recognition by complement receptors [14,15].

Abbreviations: FI, complement factor I; CpFI, factor I of *Chiloscyllium plagiosum* (whitespotted bamboo shark); LP, leader peptide; FIMAC, factor I/ membrane attack complex domain; SRCR, scavenger receptor cysteine rich domain; LDLRA, low-density lipoprotein receptor class A domains; SP, serine protease domain; RACE, rapid amplification of cDNA ends; GSP, gene-specific primers; qRT-PCR, real-time quantitative RT-PCR; LD PCR, long distance PCR; UTRs, untranslated regions; GcFI(s), factor I isotypes of *Ginglymostoma cirratum* (nurse shark); TrscFI, factor I of *Triakis scyllium* (banded houndshark); SSR, shark-specific region.

* Corresponding author. Tel.: +86 25 83271016; fax: +86 25 83271249.
E-mail address: yebp2013@163.com (B. Ye).

http://dx.doi.org/10.1016/j.fsi.2014.07.031
1050-4648/

In humans, FI is a serum glycoprotein comprising of 583 amino acids in length with a relative molecular weight of 88 kDa [16]. The protein is a heterodimer composed of disulfide-linked heavy and light chains of 50 and 38 kDa [16,17]. Each chain contains three N-linked glycosylation sites contributing 20–25% (w/w) of the apparent protein molecular mass [18,19]. Both heavy and light chains are encoded by a same gene located on chromosome 4q25, which comprises 13 exons and spans 63 kb [20–22]. Analysis of the primary structure of FI reveals a strong correlation between the exonic organization of the gene and the modular structure of the protein [23]. The protein modules are arranged linearly, which share sequence similarity with domains found in complement and non-complement proteins [24].

Orderly, the heavy chain consists of the leader peptide (LP), FI membrane attack complex domain (FIMAC), a CD5-like domain (also known as the scavenger receptor cysteine rich domain, SRCR) and two low density lipoprotein receptor class A domains (LDLRA 1 and 2); whearas the light chain of FI covers the serine protease domain (SP) [25,26], which contains the His-Asp-Ser catalytic triad [27]. However, the FIMAC and LDLRA domains are also found in complement C6 and C7 [28], while the former is a distant member of the larger follistatin superfamily that might play a role in protein–protein interactions [29]. The SRCR is an ancient and conserved domain of ~110 residues, which is involved in ligand binding and associated with the immune system [30,31]. In human, FI is comprised of 40 cysteines, 36 of them are involved in intra-domain disulfide bridges, and the remaining four cysteines (C15–C237, C309–C435) connect the FIMAC and LDLRA1 domains and the SP domain with the heavy chain [24].

So far, the complete primary structure of FI determined by cDNA sequence has been reported in human [16,20], mouse [32], rat [33], *Xenopus* [34], teleosts [35,36] sharks [37,38] and lamprey [39]. However, when compared to mammalian FI, additional sequences are present in the heavy chain of several non-mammalian species, besides the characteristic domains: in *Xenopus*, a highly charged segment has been shown in the heavy chain following the LDLRA2 [34]; as for fishes, an extended cDNA sequence between LP and FIMAC sequence has been reported in the common carp [35], channel catfish [36], houndshark [37] and nurse shark [38], which, hitherto has not described in other vertebrates. Special is that, the additional segment in sharks contains the repeat sequences of same size, which might be evolved from an ancestor of similar gene organization pattern. It has been deduced that the inserted sequences might be a reflection of the species-specificity [38].

In this study, based on the homogenous sequences from liver cDNA library constructed previously [40], a full-length of FI cDNA (CpFI) has been cloned from whitespotted bamboo shark (*C. plagiosum*), which is widely distributed along the Eastern and South China Sea. Multiple alignment has showed that the modular structure of CpFI is identical with mammalian FI except for an additional novel shark-specific region between LP and FIMAC, which is formed mainly by repeat sequences. Moreover, phylogenetic analysis and tissue distribution of CpFI has been comformed, which might facilitate our knowledge on the evolution of FI and relevant complement regulators in ancient vertebrates.

2. Materials and methods

2.1. Animal

A male whitespotted bamboo shark (*C. plagiosum*), about 0.8 kg, was obtained from Huimingqiao seafood market of Nanjing, China. The animal was transported to Marine Bioresources laboratory (China Pharmaceutical University) in a 60 L tank in aerated artificial seawater. Then it was anesthetized with chloral hydrate (700 mg/kg body weight, intraperitoneal injection), and sacrificed for meticulous dissection of tissues while avoiding cross-contamination. All dissected tissues were immediately frozen in liquid nitrogen and stored at −80 °C for nucleic acid extractions.

2.2. Cloning of the full cDNA length of CpFI

Total RNA from shark liver was extracted by TRIzol® Reagent (Invitrogen Life Technologies). First-strand cDNA ready for 5′- and 3′-RACE were respectively synthesized from 1 μg total RNA using SMARTScribeTM Reverse Transcriptase provided in a SMARTer RACE cDNA Amplification Kit (Clontech, USA).

According to the EST sequence of CpFI from our regenerative liver cDNA library [40], a series of gene-specific primers (GSPs) were designed (Table 1). These primers were used respectively in conjunction with the universal primer (supplied by Clontech) and the 3′- and 5′-RACE-ready cDNA, using Advantage 2 PCR kit (Clontech) according to the manufacturers' protocol. Amplification was performed as: 94 °C for 30 s, 68 °C for 30 s and 72 °C for 3 min for 30–35 cycles. The target products, produced by nested PCR, were gel purified and subcloned into pEasy-T3 vector (Transgen Biotech, China) for sequence determination.

As each consecutive transcript was cloned and sequenced, the overlapping compilation represented the target for full-length amplification. Finally, full coding sequence of CpFI cDNA were amplified by long distance PCR (LD-PCR) using the primers designed at the end of 5′- and 3′- untranslated regions (UTRs). All primers used in the study were listed in Table 1. The program was as follows: 94 °C for 3 min, 30 cycles of 94 °C for 30 s, 52 °C for 30 s, and 72 °C for 3 min, followed by a final extension at 72 °C for 10 min. Then the PCR products were subcloned for sequencing.

2.3. Bioinformatic and phylogenetic analysis of CpFI

The identities of full-length nucleotide sequence of CpFI were established in Basic Local Alignment Search Tool (BLAST) search engine [41], while the tandem repeats between LP and FIMAC were located by “Tandem Repeats Finder” [42]. Protein structure predictions were performed using software at the ExPASy Molecular Biology Server (http://expasy.pku.edu.cn) and Simple Modular Architecture Research Tools (http://smart.embl-heidelberg.de/) [43]. Glycosylation pattern and signal peptide were predicted at CBS Prediction Servers online (http://www.cbs.dtu.dk/services/).

Amino acid sequences of FI from various species were downloaded from the Genbank database and multiple sequence alignment was conducted using the Clustal X2 software [44]. Phylogenetic tree was constructed based on the deduced amino acid sequences using the neighbor-joining (NJ) program of MEGA version 5.1 [45].

2.4. Preparation of antisera

To examine the tissue distribution of CpFI protein, rabbit antiserum was prepared by repeated immunizations as described [46]. From immunogenicity predictions, partial protein sequence of CpFI (residues 476–671) as deduced from the cDNA, was recombinantly expressed. The cDNA fragments that encode the recombinant protein was amplified using the primers for expression vector construction (Table 1) and subcloned into pEASY-E2 vector (Transgen Biotech, China), then transformed into competent *Escherichia coli* BL21 (DE3) cells for recombinant expression. The fusion proteins were purified via His-Bind resin chromatography (Novagen, USA) according to the manufacturer's instructions. The purified protein was emulsified in an equal volume of complete Freund's adjuvant (Sigma, USA) for primary immunization followed