A THESIS FOR THE DEGREE OF MASTER OF SCIENCE

Molecular Insights into the Immunological Responses in Teleost under Pathological Conditions via Characterization of Four Complement Genes from Rock Bream (*Oplegnathus fasciatus*)

Gelshan Imarshana Godahewa

DEPARTMENT OF MARINE LIFE SCIENCES

GRADUATE SCHOOL

JEJU NATIONAL UNIVERSITY

REPUBLIC OF KOREA

February 2015

Molecular Insights into the Immunological Responses in Teleost under Pathological Conditions via Characterization of Four Complement Genes from Rock Bream (*Oplegnathus fasciatus*)

Gelshan Imarshana Godahewa

(Supervised by Professor Jehee Lee)

A thesis submitted in partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE

February 2015

The thesis has been examined and approved by

.....

Thesis Director, **Choon Bok Song**, Professor of Marine Life Sciences
School of Bio Medical Sciences, Jeju National University

.....

Bong-Soo Lim, Research Professor
Fish Vaccine Research Center, Jeju National University

.....

Jehee Lee, Professor of Marine Life Sciences
School of Bio Medical Sciences, Jeju National University

Date:.....

DEPARTMENT OF MARINE LIFE SCIENCES
GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY
REPUBLIC OF KOREA

I dedicate my dissertation to my father for being the shadow of my life, guiding me to the correct path. Without you I won't be here today as a great person

Also, this dissertation is dedicated to my fiancée who has been a great source of motivation, inspiration and backing through all these years

Acknowledgement

I would like to offer my first and foremost thanks and gratitude to my supervisor, Professor Jehee Lee, Jeju National University for his guidance, understanding, patience and most importantly his immense knowledge and professional experience during my graduate studies at the Marine Molecular Genetics Laboratory, Jeju National University, Jeju, Republic of Korea. From finding an appropriate subject in the beginning to the process of writing thesis, he offers his unreserved help and guidance and led me to finish my thesis step by step. His mentorship was paramount in providing a well-rounded experience consistent my long-term career goals. He encouraged me not only to grow as an experimentalist but also to finish this dissertation academically while guiding me through the rough road as an independent thinker. During my study period in Jeju National University, I have adored him as a great advisor, good leader with goal-oriented personality. I am really grateful to have him in my enlighten event of my life.

I would also like to thank Dr. Mahanama De Zoysa for his assistance and guidance in getting my graduate career started on the right foot and providing me with the foundation for becoming a research scholar by introducing me to Prof. Jehee Lee. I am especially thankful to Dr. Barana Jayawardana and Dr. Janak K. Vidanarachchi, who kindly offered this opportunity to me through Dr. Mahanama De Zoysa.

Especially, I would like to extent my sincere thanks and give credit those members of my masters committee, Prof. Choon Bok Song and Dr. Bong-Soo Lim for their input, valuable discussions and accessibility. Also my gratitude goes for the other professors in my department for their expertise, patience and conductive comments during my coursework.

I extend my sincere thanks to Dr. Ilson Whang and Dr. Qiang Wan for their kindness and valuable time spent on my research work and writings. I would like to acknowledge Dr. Bong-Soo Lim and Dr. Hyung-Bok Jung and other members of Marine and Environmental Research Institute of Jeju National University, who kindly supported me during field and laboratory experiments of my study.

I remember the generosity and encouragement of past and present colleagues of our laboratory and I am thankful to Dr. Kang, Dr. Saranya, Dr. Niroshana, Dr. Umasuthan, Yucheol Kim, Sukkyoung Lee, Hyowon Kim, Sanjaya Bathige, Anushka Elvitigala, Thulasitha William, Minyoung Oh, Thiunuwan Piyathilaka, Seongdo Lee, Viraj Udayantha, Lalinka Herath, Handun Eranga, Mothi Shri Seran, Eunyoung Jo, , Jiyeon Ko, Sun Hye Kang, Jeongin Ma, Taesoo Kim and Eunju Yang for their valuable help in research work as well as in numerous ways in my life in Korea who have given me their unequivocal support throughout, as always, for which my mere expression of thanks likewise does not suffice. Further, I would like to express my profound gratitude and thanks to Prasad Tharanga, Dilshara Matharage, and Chaminda Lakmal for helping me in various ways. I am also grateful to Department of Marine Life Sciences for providing me a peaceful environment for my academic and research works. I praise the enormous amount of help and teachings to all those whom I have not mentioned above but helped me in numerous ways to my success.

My warmest gratitude goes to my father, brother and family members for their love, care, moral support and advice throughout my life.

Last but not least, I want to express my enormous gratitude to my ever loving fiancée, Chathurika Nadeeshani Perera for her priceless encouragement and unflagging love gave to me at every time. This successful dissertation is simply impossible without you.

요약문

보체 시스템은 초기면역체계의 하나로서, 병원체가 침입하였을 때 항체를 경유하며, 옵소니화하기 위해 일련의 염증반응을 일으킨다. 보체 단백질들은 단백질가수분해효소로 불활성 전구체로 분비 되다가 병원체의 침입시 그들 스스로 가수분해함으로써 감염부의에 국부적으로 활성을 띄며, 강한 염증반응을 일으킨다. 보체 활성화는 병원체감염에 대항하여 3 가지 경로가 있으며, 첫번째 경로는 활성화된 보체 단백질이 침입한 세균에 공유결합을 통하여, 식세포들이 이들을 인식해 옵소니화를 일으킨다. 두번째 경로는, 보체 단백질의 작은 조각들이 (fragments) 보체 활성 부위로 식세포들을 모으고, 이들을 활성화시키기 위한 화학주화성 인자로 작용하며, 세번째 경로는 termiral 보체 구성요소들이 세포막에 구멍을 내어 침입한 병원균에 대해 방어하는 기작을 가지고 있다.

포유류에서의 보체 시스템에 관한 활성 및 기능적 특성, 다른 신호전달경로의 활성등 연구가 다양하게 되어있지만 어류에서의 보체 인자에 관한 연구는 포유류에 비하여 미미한 실정이다. 경골어류인 돌돔은 한국 양식 산업에서 주목 받고 있는 어종이며, 전염병이 돌기 쉬운 어종 중 하나이다. 그러므로 감염성 질병에 대한 어류의 방어체계를 이해하는 것이 필요하며, 이 논문에서는 감염에 따른 선천 면역체계를 분자유전학적 수준에서 유전자를 동정하였고, 이는 효과적인 질병 관리에 많은 도움을 줄 것이다.

이 연구에서는, 4 개의 필수 보체 구성성분 complement component 1r (C1r), complement component 1s (C1s), complement factor D (CfD) carboxypeptidase N1 (CPN1)을 동정하였고, 보체시스템에서 관련에 기초하여 보체 고전경로의 개시자 구성성분 (C1r 과 C1s), 대체경로의 변환자 (CfD), 보체 활성 시스템의 아나필락시톡소 불활성물질 (CPN1)에 따라서 총 3 개의 장으로 나뉜다.

제 1 장 에서는 돌돔의 C1r (OfC1r)과 C1s (OfC1s)의 전사적 수준과 분자적 및 genomic level 에서 연구 되었다. OfC1r 과 OfC1s 의 cDNA 는 각각 2533 bp 와 2724 bp 로 구성되었고, 704 개 와 691 개의 아미노산 서열로 암호화되었다. 다양한 생물학정보를 이용하여 OfC1r 과 OfC1s 모두 signal 펩타이드, 2 개의 CUB domain, 칼슘결합 EGF domain, 2 개의 보체 조절 단백질 (CCP), 세린 프로테아제 트립신 domain (SPr)과 같은 유사한 domain 을 확인하였다. OfC1r 는 *Pundamilia nyererei* C1r (73.4 %) 과 OfC1s 는 O*ryzias latipes* C1s (58.2 %)로 가장 유사했고, 계통학적 분석을 통하여 OfC1r 과 OfC1s 는 서로 다른 2 개의 그룹으로 나뉘고, 모두 어류 그룹에 속했으며, genomic 분석을 통하여 OfC1r coding region (CDS)은 13 개의 exon 과 12 개의 intron 으로 OfC1s CDS 는 11 개의 exon 과 10 개의 intron 으로 구성되었다. OfC1r 의 조직별 mRNA 발현 양상은 혈액과 간, OfC1s 는 간과 비장에서 높은 발현양상을 보였으며, 박테리아 (*Edwardsiella tarda* 와 *Streptococcus iniae*) 감염초기에 OfC1r 와 OfC1s 는 간에서 유의적인 발현양상을 보여주었다. 이 연구를 통해 OfC1r 와 OfC1s 는 돌돔에서의 세균성 감염에 선천 면역 반응으로써 초기 방어 기작으로 역할을 할 것이라고 생각된다.

제 2 장은 Of*CfD* 의 분자적 characteristics 와 transcriptional properties 관한 내용이다. CfD 는 대체 경로 (Alternative pathway)에서 C3 convertase (C3bBb)을 만드는 중요한 역할을 하고 있는 유전자 중 하나이다. Complement cDNA OfCfD 는 1100bp 이며, 이 중 CDS 는 239 개의 아미노산으로 구성하는 720 bp 으로 되어있다. Characteristic serine protease trypsin superfamily domain 과 세 개의 active sites (^{31}H, ^{78}D, ^{183}S), 세 개의 substrate binding sites (^{177}D, ^{198}S, ^{201}G)를 mature protein 에서 찾을 수 있었으며, Nile tilapia (*Oreochromis niloticus*)의 CfD 에서 identity (73.8 %) 와 similarity (78.9 %)를 보였다. Multiple sequence alignment 는 trypsin superfamily domain 이 CfD homologue 에서도 잘 보존되었음을 알 수 있고, 계통분류학적 분석결과 이 연구에서 확인하고자 하는 다른 어류들과 같은 그룹을 형성하고, OfCfD 는 넙치와 가장 가까웠다. 3 차원 구조 모형화를

통하여 인간의 CfD X-ray crystal 와 같은 구조를 형성함을 알 수 있었다. OfCfD 의 Genomic DNA 는 2451 bp 로 5 개의 exon 과 4 개의 intron 으로 구성되어 있다. 이 유전자는 비장과 간에서 제일 높은 발현을 보았고, *S. iniae*, lipopolysaccharide (LPS), rock bream iridovirus (RBIV) 와 Poly: IC 면역자극을 주었을 때 비장에서 유의적으로 발현이 증가함을 보였다. 이 장에서는 돌돔에서 병원체가 감염되었을 때, Of*CfD* 는 대체경로에서 중요한 초기 방어 기작을 담당하는 것으로 생각 된다.

제 3 장에서는 분자 기능, 게놈 배열과 anaphylatoxin 의 inactiviator 으로 보체 cascade 에서 OfCPN1 의 면역 반응 전사 조절에 대해서 설명하려고 한다. OfCPN1 는 The full-length cDNA 서열 1591 bp 과 CDS 는 1347 bp, 488 개의 아미노산을 가지고 있으며, signal peptide, Zn-binding region 1 signature, Zn-binding region 2 signature, peptidase-M14 Zn carboxypeptidase site, Zn-binding site 1, Zn- binding site 2 와 N-glycosylation site 도메인들을 찾을 수 있었다. 계통 및 pairwise homology sequence 는 Dicentrarchus labrax 와 가장 높은 identity 를 보였으며, Multiple sequence 분석을 통하여 Zn binding sites 가 다른 CPN1 homologue 에서 높게 보존되었음을 확인하였다. *OfCPN1* 의 게놈 배열은 9 개의 exon 과 8 개의 intron 으로 구성됐으며, qPCR 분석을 통하여 가장 높은 발현은 간, 두번째는 신장에서 높은 발현양상이 높았고 돌돔에 면역자극 (*E. tarda, S. iniae*, LPS, RBIV 와 Poly: IC) 을 주었을 때 간 조직에서의 mRNA 의 발현이 up-regulation 되었다.

종합적으로, 이 연구는 모델 생물 돌돔 (*Oplegnathus fasciatus*)을 이용하여, 4 개의 보체 유전자를 분자적 관점에서 연구하였으며 더 나아가 이는 선천 면역 메커니즘의 기존 지식의 확장에 많은 기여를 제공할 것이다. 또한 이러한 결과들은 경골어류 보체 시스템에 대한 이해를 도울 수 있을 것이라고 생각된다.

Summary

The complement system is a group of plasma proteins that reacts with one another to opsonize pathogens via antibodies and induce a series of inflammatory responses as part of the innate immune system. Among them number of complement proteins are proteases that are activated by proteolytic cleavage by themselves. Such enzymes are called zymogens and are stored inside the cells and secreted as inactive precursor enzymes. In the complement system, the precursor zymogens are widely distributed throughout the body fluids and tissues to response for the pathogenic infection. However, they are activated locally and trigger series of potent inflammatory events at the sites of infection. The complement system activates through a triggered-enzyme cascade. In such a cascade, an active complement enzyme is generated by cleavage of its zymogen precursor then cleaved its substrate, another complement zymogen, to its active enzymatic form. The activation of an initiator complement proteins at the start of the pathway is hugely amplified by each successive enzymatic reaction, resulting in the rapid generation of a disproportionately large complement response. As might be expected, there are many regulatory mechanisms to prevent uncontrolled complement activation via the function of complement anaphylatoxin inactivators and complement inhibitors.

There are three distinct pathways through which complement can be activated on pathogen surfaces. These pathways depend on different molecules for their initiation, and there are three ways in which the complement system protects against infection. First, it generates large numbers of activated complement proteins that bind covalently to pathogens, opsonizing them for engulfment by phagocytes bearing receptors for complement. Second, the small fragments of some complement proteins act as chemoattractants to recruit more phagocytes to the site of complement

activation, and also to activate these phagocytes. Third, the terminal complement components damage certain bacteria by creating pores in the bacterial membrane.

In mammalians, complement system is well studied in relation to the activations, direct functional characteristics, activation of other signaling pathways, drawback of the needless activations etc. However, complement components have been poorly studied in teleost fish or identified arbitrarily in several fish species. Therefore, immunological study of complement components from one of the economically important teleost fish species rock bream, is a valuable investigation for molecular immunologists. Rock bream (*Oplegnathus fasciatus*), is demanding aquacrop in South Korea which is prone to infectious diseases. The studies on immunological perspectives of fish and genetic background are important to gaining a detailed understanding of fish immune system and may immensely help in development of the effective disease control strategies. Understanding of molecular insights into mRNA expression of vital complement components of the complement cascade system is enormously relief, to evaluate the first line of defense mechanism. Thus, scientific studies will be able to obtained clear illustration of innate immune function of complement components and will be able to develop novel strategies to prevent diseases of rock bream fish farming in South Korea.

In this study, four vital complement genes including the complement component 1r (C1r), complement component 1s (C1s), complement factor D (CfD) and carboxypeptidase N1 (CPN1) have been identified and characterized at molecular, genomic and transcriptional level as an initiator, convertor and terminator complement components, respectively. This report is divided into three main chapters based on its involvement in complement system, such as initiative components of the

classical pathway (C1r and C1s), convertor molecule of the alternative pathway (CfD) and anaphylatoxin inactivator of the complement cascade system (CPN1).

In chapter I, rock bream C1r (OfC1r) and C1s (OfC1s) from classical pathway were characterized at molecular and genomic level while analyzing their transcriptional modulation under pathological conditions. The complete cDNA sequences of OfC1r and OfC1s were consisted of 2533 bp and 2724 bp, respectively. Their putative polypeptide sequences were 704 aa (78 kDa) and 691 aa (76 kDa), respectively. Subsequently, the identified sequences were characterized using various bioinformatics tools, while comparing the sequences with other known similitudes. Both OfC1r and OfC1s shared similar domain architecture including a signal peptide, two CUB domains, a calcium binding EGF domain, two complement control protein (CCP) domains and serine protease trypsin domain (SPr). OfC1r and OfC1s shared highest identity with *Pundamilia nyererei* C1r (73.4 %) and *Oryzias latipes* C1s (58.2 %), respectively. Phylogenetic analysis revealed that both OfC1r and OfC1s were tightly clustered with the fish clade while they were grouped into two different main clusters. According to the genomic organization, the OfC1r coding region (CDS) was distributed across thirteen exons interrupted by twelve introns whereas; OfC1s CDS was distributed within eleven exons interrupted by ten introns. The mRNA of OfC1r was significantly expressed in blood and second most in liver. Though, the mRNA of OfC1s was highly expressed in the liver and spleen, as indicated by quantitative real-time PCR (qPCR) tissue profiling. In addition, both OfC1r and OfC1s were transcriptionally responded to immune challenges, with liver expression being significantly up-regulated in the early phase of infection with live bacteria (*Edwardsiella tarda* and *Streptococcus iniae*). These data collectively suggest that the

OfC1r and OfC1s may play first line defense roles as an innate immune response to protect the rock bream from bacterial infections.

Chapter II enlightened the molecular characteristics and transcriptional properties of OfCfD from alternative pathway. The CfD is one of the key molecules in alternative pathway which makes C3 convertase (C3bBb). The complete cDNA sequence of OfCfD consisted of 1100 bp with 720 bp of CDS, encoding polypeptide of 239 aa (26 kDa). Characteristic serine protease trypsin superfamily domain, three active sites (^{31}H, ^{78}D, ^{183}S) and three substrate binding sites (^{177}D, ^{198}S, ^{201}G) were identified in mature protein of OfCfD. CfD orthologue of *Oreochromis niloticus* exhibited the maximum identity (73.8 %) and similarity (78.9 %) with OfCfD. According to the multiple sequence alignment, CfD homologues showed the conservation in the trypsin superfamily domain. Phylogenetic analysis of OfCfD evidenced to its higher evolutionary proximity with orthologue from olive flounder, among the other fish similitudes. The modeled tertiary structure of OfCfD resembled the characteristic features of the folding of human CfD X-ray crystal structure. The complete genomic DNA of OfCfD was 2451 bp, consists of five exons interrupted by four introns. In spatial mRNA expression profile, OfCfD transcripts were constitutively expressed in all the tissues while highest was observed in spleen and second most in liver. In challenge experiments; OfCfD transcripts were significantly up-regulated in spleen tissue post challenge with *S. iniae***,** lipopolysaccharide (LPS), rock bream iridovirus (RBIV) and Poly: IC. Collective findings of this chapter may suggest that the OfCfD crucial in first line defense mechanism involved in alternative pathway of the rock bream upon pathogenic infections.

Third chapter describes the molecular features, genomic arrangements and immune responsive transcriptional regulation of OfCPN1 from the complement

cascade as an anaphylatoxin inactivator. The full-length cDNA sequence was 1591 bp which possessed the 1347 bp CDS. The predicted molecular mass of putative polypeptide of OfCPN1 (488 aa) was 51 kDa. The domain architecture evidenced a signal peptide, Zn-binding region 1 signature, Zn-binding region 2 signature, peptidase-M14 Zn carboxypeptidase site, Zn-binding site 1, Zn- binding site 2 and N-glycosylation site. According to the phylogenetic and pairwise homology sequence comparison, OfCPN1 showed highest identity to *Dicentrarchus labrax*. Multiple sequence analysis revealed the highly conserved Zn binding sites at the Zn binding regions among the CPN1 homologues. Genomic arrangement discovered that OfCPN1 consisted of nine exons interrupted by eight introns. The qPCR analysis confirmed that OfCPN1 gene was constitutively expressed in all examined tissues, isolated from healthy rock bream, and the highest expression was observed in liver and second most in kidney. In temporal mRNA expression of OfCPN1 from immune challenged rock bream, significant up-regulations were detected in liver tissues upon *E. tarda, S. iniae,* LPS, RBIV and Poly: IC challenges. Taken together, the molecular characteristics and transcriptional responses to immune stimulants suggest that OfCPN1 is involved in immune responses upon the complement system activation in of rock bream.

Collectively, this study perceives molecular perspectives of four complement genes of rock bream fish *Oplegnathus fasciatus* using as a model organism. Moreover, this furnished a substantial contribution to the extension of existing knowledge in innate immune mechanism and it will be important to understand their role in relation to survival of fish from challenging environment with abundant pathogenic threats. Further, these findings make significant understanding into the immunological perspectives of the teleost complement system.

Table of Content

List of Figures

List of Tables

1.0. Introduction

1.1. General introduction

This comprehensive study is to explore the immunological function and molecular aspects existence in four complement proteins from rock bream, *Oplegnathus fasciatus.* The complement system is a crucial part of the innate immune system which helps antibodies and phagocytic cells to clear pathogens from an organism. In order to achieve the immune function, the complement system comprises of 35 distinct proteins scattered in biota. The direct functional roles and activation, activation of downstream signaling pathways, drawback of the needless activations of this system have been extensively investigated in mammals. However, the transcriptional and functional characterization of complement proteins is comparatively lesser in fish species. In this study, with the advances made in genomics and transcriptomics, four vital complement genes including, complement component 1r, complement component 1s, complement factor D and anaphylatoxin inactivator (carboxypeptidase N) have identified and characterized at molecular, genomic and transcriptional level. Interestingly, understanding of the crucial complement genes as an initiator, convertor and terminator of the complement system could help us in enlightening the knowledge of innate immune system in fish. According to this regard, investigation of immune mechanisms, functioning in rock bream is an effective approach to invent the appropriate pathogenic disease preventive schemes. Since innate immune system is the fundamental defense mechanism in fish, which stimulate the adaptive immunity, identification and revealing the roles of innate immune parameters will be an inauguration in elucidating the different strategies of fish defense system (Magnadottir, 2006).

1.2. Rock bream as an aquacrop in South Korea

Mariculture is a subdivision of aquaculture, which is dedicated to cultivate marine organisms for human consumption. Beneficially, the Korean peninsula is surrounded by the sea except on the north. Most marine fish farms are situated in the southern coastal area of the peninsula, followed by western and eastern coastal areas. Currently, aquaculture production in South Korea is dominated by seaweeds, mollusks and finfish. Among those mariculture creatures, fish are used as the main delicacy in most part of the world, prominently in South Eastern Asian countries like Korea, China and Japan. Rock bream is a fish species of knifejaw also known as barred knifejaw and striped beackfish, an economically important marine aquacrop harvested by commercial fisheries and mariculture farming inhabited in the coastal areas of the Pacific and Indian oceans. However, coastal waters of East Asia, is highly prone to deadly infectious diseases caused by various pathogens (Li et al., 2011; Park Ii, 2009), due to infections of bacteria like *Edwardsiella tarda, Streptococcus iniae* and virus like rock bream iridovirus (RBIV). Those evidences affirming the necessity of a precise disease management strategy in rock bream mariculture farming for its sustainability (Park Ii, 2009). Hence, search for the molecular information about the immune mechanisms of this species might strengthen the endeavors undertaken to control or prevent its mortality. Therefore, in an attempt to explore the genes involved in the innate immunity of rock bream against potential bacterial and viral pathogens using transcriptomic approach, the complement genes were targeted and identified.

Figure 1.1. Rock bream fish *Oplegnathus fasciatus*

1.3. Fish immune system

The vertebrate immune system consists of two arms namely a specific adaptive and a non-specific innate system. Evolutionarily, fish being the preliminary vertebrates possesses both the arms of immunity. Adaptive immunity first appeared about 450 million years ago in cartilaginous and bony fishes. It plays an important role in the protection against recurrent infections by generating 'memory' cells and specific soluble and membrane-bound receptors, such as immunoglobulins (Ig) and T cell receptors (TCR), which allow for the fast and efficient elimination of the specific pathogens. The development of vaccines as one of the means of controlling infection relies on the principle of adaptive immunity, and vaccination technologies have revolutionized the fish farming industry, as they previously have domestic animal and human medicine. The major difference of the fish system with other vertebrates is that their metabolism and immune response is temperature dependent. Therefore, development of vaccines for fish diseases was recognized as a primary requirement of disease control in aquaculture. Vaccine development in fish culture has reached a new era due to innovation of novel biotechnological tools and use of modern molecular biological methods such as gene silencing, RNA interference, micro RNA, and microarray. Identification and characterization of immune related genes is one of the fundamental requirements to understand the immune system and their role in the facilitation of functions of vaccine delivery.

1.3.1. Innate immunity

Innate immunity is an evolutionarily conserved, non-specific, first line of defense genetically inherited and not learned, adapted or permanently heightened, because of exposure to microorganisms or vaccination. The innate immune system

comprises of the cells and mechanisms that defend the host during microbe exposure and initial adaptive responses. The innate immune system recognizes microbes through pattern recognition receptors (PRRs), which are receptors specific for molecular components of microorganisms that are not relevant for the host. Some innate immune responses are temporarily up-regulated as a result of exposure to microbes, but the components of the innate immune system do not change permanently during an individual's lifetime. Unlike the adaptive immune system, it does not confer long-lasting or protective immunity to the host (Fearon and Locksley, 1996). Innate immune system is present in all classes of life ranging from microorganisms to well-developed mammals.

1.3.2. Adaptive immunity

The adaptive immune system is composed of highly specialized, systemic cells and processes that eliminate or prevent pathogenic growth. Thought to have arisen in the first jawed vertebrates, the adaptive or "specific" immune system is activating by the "non-specific" and evolutionarily older innate immune system (which is the major system of host defense against pathogens in nearly all other living things). The adaptive immune response provides the vertebrate immune system with the ability to recognize and remember specific pathogens (to generate immunity), and to mount stronger attacks each time the pathogen is encountered. It is adaptive immunity because the body's immune system prepares itself for future challenges.

1.4. The complement system

The complement system was originally discovered in 1890s when it was found to aid or "complement" the killing of bacteria by heat-stable antibodies present in normal serum (Walport, 2001). The complement system is composed of around 35 proteins, which are found as soluble proteins circulating in blood or as membrane-associated proteins. These proteins play key roles in complement-mediated killing of pathogens through the lytic pathway (Gal et al., 2007; Korotaevskiy et al., 2009). The complement system exerts humoral effects in innate immunity and plays a crucial role in recognition, tagging and clearance of invading microorganisms (Nakao et al., 2011). It acts as a bridge in between innate and adaptive immune system. The activation of the complement system leads to a sequential cascade of enzymatic reactions resulting in the formation of the potent anaphylatoxins C3a and C5a that elicit a plethora of physiological responses that range from chemoattraction to apoptosis. Three distinct pathways are responsible for target recognition and formation of a protease complex for C3 activation: the classical, lectin and alternative pathways congregate into a final common pathway, known as the cytolytic pathway. This cytolytic pathway is responsible for formation of the membrane attack complex (MAC) that ultimately causes lysis of the targeted foreign cell (Biesecker and Muller-Eberhard, 1980). The complement proteins exist as inactive zymogens under normal circumstances, when sequentially cleaved by proteases become activated. On the other hand, some components of the system serve as the substrate of an earlier component of the cascade and then as an enzyme to activate a subsequent component. As a whole, there are four main functions of the complement system; lysis of microorganisms, opsonization, chemotaxis and agglutination (Boshra et al., 2006).

1.4.1. Classical pathway

The classical pathway is an activation of complement system via antibody responses. A wide variety of pathogenic microorganisms, including a number of bacteria and viruses, as well as many infected cells, efficiently activates the classical pathway after their recognition by antibodies. The primarily classical pathway is initiated by the C1 complex, which is circulating in the blood serum. The C1 complex consists of six molecules of C1q, two molecules of C1r and C1s. Activation is triggered upon interaction of the serum C1 complex with antigen–antibody complexes or immune aggregates containing immunoglobulin G (IgG) or IgM (Wimmers et al., 2003). C1 binding to immunoglobulins is mediated by its C1q subunit and involves the heterotrimeric C terminal globular 'heads' of the protein. Subsequently, C1s and C1r get activated and produce an enzyme that cleaves two complement components (C4 and C2) and continue cascade system to assemble of the C3 convertase.

1.4.2. Lectin pathway

A second activation route and a type of cascade reaction of the complement system were discovered in the early 1980s. In contrast to the classical complement pathway, the lectin pathway does not recognize an antibody bound to its target. It is triggered by a C1-like complex protease in which the recognition function is mediated by mannan-binding lectin (MBL), a member of the collectin family (Malhotra et al., 1994). Unlike other collectins, MBL is able to mediate the opsonization of microorganisms through its multiple C-terminal C-type lectin modules. MBL identifies ranges of oligosaccharides present in the surface of wide spectrum of microorganisms, including yeasts (*Candida albicans*), viruses (HIV, influenza A), and

a number of bacteria (*Salmonella montevideo*, *Listeria monocytogenes*, *Haemophilus influenzae* and *Neisseria meningitidis*) in human (Arlaud and Colomb, 2001).

1.4.3. Alternative pathway

The third subdivision of the complement system is an innate component of the immune system, upon natural defense against pathogenic infections. It is initiated by the spontaneous hydrolysis of C3, which is abundant in the blood plasma. The activation of the C3 involves complement factor D (CfD) coupled with Factor B and Factor P. Then CfD splits the bound Factor B into Bb and Ba, to form the C3(H_2O)Bb enzymatic molecule. Afterward, alternative C3 convertase is synthesized to continue the MAC formation and cell lysis.

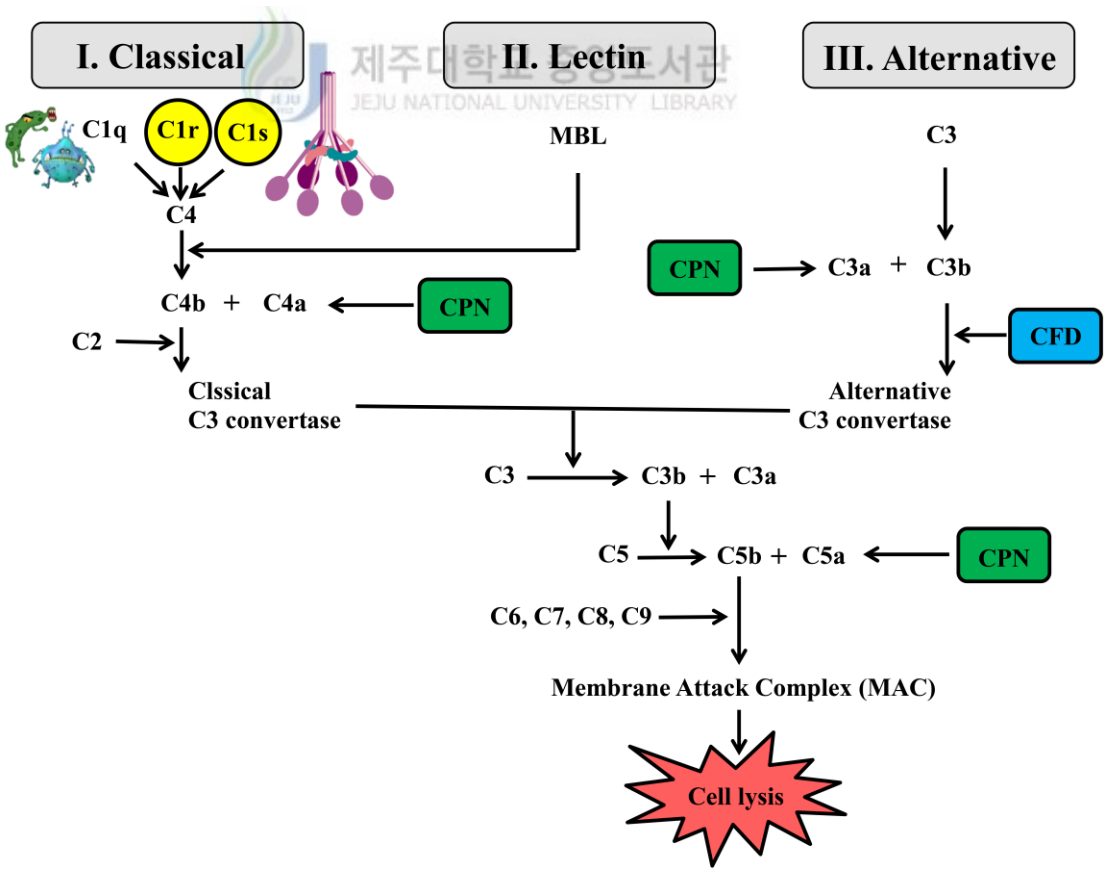


Figure 1.2. Complement system

1.5. Objectives of the study

The main focus of this study was to discover the molecular evidences in to immune mechanism of four complement genes from rock bream (*Oplegnathus fasciatus*). This dissertation included the two complement genes from classical pathway as initiators; complement component 1r (*OfC1r*) and complement component 1s (*OfC1s*), one complement gene from alternative pathway as stimulator; complement factor D (*OfCfD*) and complement carboxypeptidase N as anaphylatoxin inactivator (*OfCPN1*). The *in silico* molecular characterization, genomic structural arrangement, immune responsive transcriptional regulations and their protease activities were determined from above mentioned four complement genes.

2.0. Materials and Methods

2.1. Experimental animal, pathogens and chemicals

Healthy rock bream fish with an average body weight of 50 g, obtained from the Jeju Special Self-Governing Province Ocean and Fisheries Research Institute (Jeju, Republic of Korea) were adapted to the laboratory conditions (salinity 34 ± 1‰, pH 7.6 ± 0.5 at 24 ± 1 ℃) in 400 L aquaria in Marine Molecular Genetics Lab, Jeju National University (Jeju, Republic of Korea). Animals were accustomed to the laboratory conditions for one week prior to any experimentation. After adaptation, to minimize stress factors, a maximum of 40 animals per 40 L were housed in each tank. Rock bream iridovirus (RBIV) was isolated from the infected rock bream kidney samples. Moreover, two pathogenic bacterial strains *Edwardsiella tarda* and *Streptococcus iniae* strain were obtained from Dept. of Aqualife medicine, Chonnam National University (Korea). Rock bream iridovirus (RBIV) was obtained from the kidney of an RBIV-infected moribund rock bream. Primers used in this study were synthesized by Integrated DNA Technologies, Inc, USA. All of the chemical used in this study were in molecular biology grade and were purchased from Sigma, USA. SYBR Ex Taq and molecular markers were purchased from Takara, Japan.

2.2. Rock bream transcriptome library construction and cDNA identification

We have constructed a rock bream transcriptome sequence database by the GS-FLX™ genome sequencing technique as previously described (Droege and Hill, 2008). Briefly, the total RNA was extracted from pool of several tissues including gills, brain, blood, liver, spleen, pituitary gland, head kidney and kidney of three healthy rock bream fish using the Tri Reagent™ (Sigma, USA). The concentration and purity of RNA was determined by absorbance at 260 and 280 nm using a UV-

spectrophotometer (BioRad, USA). Then, the mRNA was purified using an mRNA isolation kit (FastTrack™ 2.0, Invitrogen, USA). First strand cDNA synthesis was carried out with the Creator™ SMART™ cDNA library construction kit (Clontech, USA). Over-representation of the most commonly expressed transcripts was excluded by normalizing the synthesized cDNA using Trimmer-Direct cDNA normalization kit (Evrogen, Russia). Thereafter, the sequencing of rock bream cDNA was performed on a GS-FLX Titanium instrument (DNA Link, Inc.). A single full-plate run was performed using the above normalized cDNAs and the reads obtained were processed and assembled with Arachne assembly program (Batzoglou et al., 2002). Finally, a cDNA GS-FLX shotgun library was created based on this sequencing data. A single putative cDNA, homologous to known complement component 1r, complement component 1s, complement factor D and carboxypeptidase N1 were identified by homology screening using the basic local alignment sequencing tool (BLAST; http://www.blast.ncbi.nlm.nih.gov/Blast), and was designated as OfC1r, OfC1s, OfCfD and OfCPN1, respectively.

2.3. Rock bream bacterial artificial chromosome library (BAC) construction

A bacterial artificial chromosome (BAC) library was custom constructed for rock bream from the isolated blood cells (Lucigen® Corp., Middleton, WI, USA). Briefly, genomic DNA obtained from peripheral blood cells was randomly sheared, and the blunt ends of large inserts (>100 kb) were cloned into pSMART BAC vector to obtain an unbiased, full coverage library. The corresponding constructs were then transformed to *E. coli* and around 92160 clones, possessing an average insert size of 110 kb (100-120 kb), were arrayed in 240 microtiter plates with 384 wells. The BAC library was stored in -80 ℃ (Quiniou et al., 2003). Aliquots of the grown cultures

were pooled with other clones from the same plate, row or column pools for DNA preparation. Then, they were further combined to form Super Pools (n = 20). For the screening purpose, the DNA was extracted from each clone separately and arranged by super pooling and pooling system, comprising 20 super pools, 16 row pools and 24 column pools, spanning the entire rock bream genome.

2.4. BAC library screening and BAC clone sequencing

Primers for BAC library screening were designed based on the cDNA sequence identified from the rock bream cDNA database (Table 2.1). A two-step PCR-based screening of the BAC library was used to identify the super pools and the individual plate, row, and columns (P, R, C) through which the location of the clone of interest bearing the respective candidate genes could be located. Each Super Pool had a corresponding 96-well plate containing its Plate-pools, Row-pools, and Column pools (P-R-C pools). The Super Pools were screened during the first round of PCR, and plates, rows, and columns were screened during the second round of PCR to determine the exact well containing the clone of interest. Finally, the identified clone was isolated from the corresponding well, and confirmed by a colony PCR with gene-specific primers. The positive BAC clones were cultured and purified with a Qiagen Plasmid Midi Kit (Hilden, Germany). Finally, 15 positive clones were pooled and subjected to Pyro-sequencing (GS-FLX 454, Macrogen, Republic of Korea).

Table 2.1 Description of primers used in this study

Gene Name	Amplification and Screening	Primer Sequence (5' - 3')
OfC1r-F	qPCR and BAC	ATGTGCCAAAGCAATGGACAGTGG
OfC1r-R	qPCR and BAC	ACTGATTCTGAAAGCCGGACAGGA
OfC1s-F	qPCR and BAC	GTTGTCATGGACAGTGAGAG
OfC1s-R	qPCR and BAC	GCCTAAATTCACCCTGGAAG
OfCfD-F	qPCR and BAC	CAACGAGGTGGTCGTTGAAGTGTTC
OfCfD-R	qPCR and BAC	CCAGAGTCACCATCACAACTGTCTTCT
OfCfD-F	ORF amplification	GAGAGAgaattcATGGCCTCCATCCAGGTGC
OfCfD-R	ORF amplification	GAGAGAaagcttTTAACTGCTCAGGTCCTGTGTCGC
OfCPN1-F	qPCR and BAC	ACGTACTTGAACTCCGGCTCCAAT
OfCPN1-R	qPCR and BAC	ACATCACGCGCATTTACAGCATCG
Ofβ-actin-F	qPCR	TCATCACCATCGGCAATGAGAGGT
Ofβ-actin-R	qPCR	TGATGCTGTTGTAGGTGGTCTCGT

F; forward, R; reverse and BAC; Bacterial Artificial Chromosome

2.5. *In silico* characterization

The complete cDNA sequences of identified OfC1r, OfC1s, OfCfD and OfCPN1 were used to analyze the putative coding sequence (CDS) and were derived its corresponding protein sequence using the DNAssist version 2.2. Functional domains and motifs were identified using the SMART proteomics database (Letunic et al., 2009), Motif Scan tools, PROSITE profile database (de Castro et al., 2006) and Conserved Domain Database (CDD) at the NCBI (http://www.ncbi.nlm.nih.gov/cdd). The putative amino acid sequences were comparatively analyzed with orthologous sequences by MatGat program to assess the identity and similarity percentages. ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) program was engaged in performing the multiple sequence alignment. The three-dimensional structures were predicted using SWISS-MODEL protein modeling server (http://swissmodel.expasy.org/). The predicted structure image was generated utilizing PyMOL v1.5 software. The phylogenetic position of OfC1r, OfC1s, OfCfD and OfCPN1 were assessed by reconstructing a phylogenetic tree using MEGA v 5.0

program (Tamura et al., 2011) based on neighbor-joining (NJ) method. To derive the confidence value for the phylogenetic analysis, bootstrap trials were replicated 5000 times.

Sequencing results of the respective BAC clones were used to determine the exon intron structures of each gene. The exon intron structures were derived by aligning the cDNA sequence with genomic sequence using Spidey program (http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey/). The genomic structures of the other species were obtained from NCBI and Ensembl genomic browser (http://asia.ensembl.org/index.htmL). The structures were constructed using Gene Mapper v2.5 (http://genemapper.googlepages.com) in order to compare the homologous genomic structures.

To discover the genomic synteny of OfC1r and OfC1s, synteny maps for the genomic neighborhoods adjacent to the C1r and C1s genes in tetrodon, medaka, zebrafish, spotted gar, chicken, frog, rat, cattle and human were obtained using the ensemble and the Genomicus genome browser (v62.02; http://www.dyogen.ens.fr/genomicus-62.02/cgi-bin/search.pl).

2.6. Spatial mRNA expression analysis

In order to profile the tissue distribution of selected complement transcripts, whole blood was withdrawn from the caudal fin of three healthy, unchallenged juvenile rock breams (~1 mL $fish^{-1}$) and centrifuged immediately for 10 min at 3000 × g at 4 ℃, to separate the peripheral blood cells (PBCs) from plasma. Subsequently, fish were killed and various tissues (muscle, intestine, brain, skin, gill, liver, spleen, heart, kidney, and head kidney) were excised and immediately snap-frozen in liquid nitrogen and stored at −80°C until total RNA extraction.

2.7. Immune challenge experiment

To study the immune response of OfC1r, OfC1s, OfCfD and OfCPN1 in different tissues, fish were divided into 6 groups (LPS, Poly:IC, *Edwardsiella tarda*, *Streptococcus iniae,* RBIV and PBS). Briefly, purified LPS (055:B5 from *Escherichia coli*; Sigma; 1.25 μg/μl), polyinosinic-polycytidylic acid (Poly I:C; Sigma; 1.5 μg/μl) were used as immune stimulant. Moreover, two pathogenic bacterial strains *Edwardsiella tarda* (Gram negative; 5×10^3 CFU/μl) and *Streptococcus iniae* (Gram positive; 1×10^5 CFU/μl) and a live virus (RBIV; 10^2 $TCID_{50}$ per fish) were used in time-course experiments.

The bacterial strains were grown in brain heart infusion (BHI) agar (Eiken Chemical Co., Japan) + 1.5 % NaCl plates at 30 ℃ for overnight. A single colony was picked into incubated at BHI broth supplemented with 1.5 % NaCl and allowed to grow at 30 ℃ with 200 rpm shaking incubator until the OD_{600} reached the absorbance of 1.0-1.5. Then, the bacterial suspension was briefly spun down at 3500 rpm for 20 min. Finally, bacterial pellet(s) was re-suspended in PBS and cell numbers were adjusted to 5×10^3 CFU/mL (*E. tarda*), and 1×10^5 CFU/mL (*S. iniae*) by serial dilution, and used as stock for experimental injection.

RBIV was obtained from the kidney of an RBIV-infected moribund rock bream. The viral specimen was prepared by homogenizing the kidney of RBIV-infected rock bream with approximately 20 volumes of 1 × PBS (1 mL PBS/100 mg tissue). The homogenate was centrifuged at 3000×g for 10 min, and the supernatant containing the virus was filtered through a 0.45 μm syringe filter, and stored at -80°C until use. For the RBIV challenge experiment, the supernatant was thawed and diluted to obtain 1×10^2 $TCID_{50}$ of RBIV per 100 μL. All the stimulants and pathogens were prepared with PBS and injected intraperitoneally (except RBIV which is administered

intramuscularly) in a total volume of 100 µl. From challenged fish, liver and spleen tissues were collected from three randomly selected fish (n=3), from each of the un-challenged, immune-challenged and PBS-injected groups at 3, 6, 12, 24, and 48 h post infections (p.i). All the tissues were flash-frozen in liquid nitrogen, and stored at −80 °C before being processed.

2.8. Total RNA extraction and first strand cDNA synthesis

Total RNA was extracted from a pool of tissue samples (n=3) by TRI Reagent™ (Sigma–Aldrich). The RNA concentration and purity was determined spectrophotometrically, by measuring the absorbance at 260 nm and 280 nm on a UV spectrophotometer (Bio-Rad, USA). Purified RNA was diluted to 1 µg/µL, and an aliquot of 2.5 µg was used to synthesize cDNA from each tissue with the PrimeScript™ first strand cDNA synthesis kit (TaKaRa, Japan). Concisely, the RNA was incubated with 1 µL of 50 µM oligo$(dT)_{20}$ and 1 µL of 10 mM dNTPs for 5 min at 65 °C. After incubation, 4 µL of 5× PrimeScript™ buffer, 0.5 µL of RNase inhibitor (20 U), 1 µL of PrimeScript™ RTase (200 U) from M-MLV were added, and the mixture was incubated for 1 h at 42 ℃. The reaction was terminated by adjusting the temperature to 70 ℃, and incubating for an additional 15 min. Finally, the synthesized cDNA was diluted 40-fold (total volume of 800 µL) and stored at −20 °C until use.

2.9. Quantitative real-time PCR (qPCR) analysis

To detect the transcript level in normal animals and immune challenged animals, the gene specific primers (Table 2.1) were used to amplify the particular transcript. qPCR was carried out using a Thermal Cycler Dice™ TP800 (Takara) in a

10 µl reaction volume containing 3 µl of diluted cDNA template, 5 µl of 2× TaKaRa Ex Taq™ SYBR premix, 0.4 µl of each of the forward and reverse primer (10 pmol/µl) and 1.2 µl of H_2O. The thermal cycling profiles included one cycle of 95°C for 30 s, followed by 40 cycles of 95°C for 5 s, 58°C for 10 s and 72°C for 20 s. Finally, a single cycle of 95°C for 15 s, 60°C for 30 s and 95°C for 15 s was executed to examine the melting curve in order to evaluate the specificity of the target amplification. Each assay was conducted in triplicates to increase the credibility. The *Livak* $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) was used to calculate the relative expression. Rock bream *β-actin* (Accession No. FJ975145) was used as the internal control gene. For the expression profiles in tissues of immune-challenged fish, the β-actin normalized mRNA levels were compared with the levels detected in the respective PBS-injected controls to calculate the fold-difference for each tissue.

2.10. Preparation of recombinant expression vector (pMAL-c2X)

To purify the recombinant protein, the cDNA fragment encoding the mature peptide of particular gene was amplified with gene specific cloning primers (Table 2.1). The primers were contained EcoRI and HindIII as two restriction sites and amplified OfCfD was cloned into pMAL-c2X vector (New England Biolabs, USA). Briefly, PCR reaction was accomplished in a 50 µL volume containing 50 ng of liver cDNA, 5 µL of 10× Ex Taq Buffer, 4 µL of 2.5 mM dNTPs, 40 pmol of each primer and 5 U of Ex Taq polymerase (TaKaRa). The reaction was carried out an initial denaturation at 94 ℃ for 3 min; 30 cycles of amplification at 94 ℃ for 30 sec, 58 ℃ for 30 sec and 72 ℃ for 1 min; and a final extension at 72 ℃ for 5 min. The resulted PCR product was distinguished on 1% agarose gel and purified using Accuprep™ gel purification kit (Bioneer, Korea). Restriction enzyme digested cDNA fragment and

pMAL-c2X expression vector (150 ng) were ligated using Mighty Mix DNA Ligation Kit (TaKaRa, Japan) at 4 ℃ overnight. The recombinant construct was transformed into DH5α competent cells. Sequence was verified (Macrogen, Korea) and was transformed into *Escherichia coli* BL21 (DE3) competent cells (Novagen) for protein expression.

2.11. Expression and purification of recombinant fusion protein

The rOfCfD-MBP fusion protein expression and purification was carried out according to the instructions of the pMAL Protein Fusion and Purification System (New England Biolabs). Briefly, BL21 (DE3) transformed cells were cultured in 500 mL LB, ampicillin (100 μg mL^{-1}) broth and 100 mM glucose at 37 ℃ until ~ 0.6 OD_{600} reached. Cells were induced with isopropyl-β-thiogalactopyranoside (IPTG) at a final concentration of 0.5 mM and the culture was incubated for 10 h at 20 ℃. Then, cells were kept on ice for 30 min and pelleted by centrifugation (4000 x g for 20 min at 4 ℃). Pelleted cells were resuspended in column buffer (20 mM Tris-HCl, pH 7.4, 200 mM NaCl) and overnight stored at -20 ℃. Cell suspension was thawed and sonicated on ice in the presence of lysozyme (1 mg mL^{-1}). The lysate was subjected to centrifugation (20,000 × g for 20 min at 4 ℃) and the supernatant was allowed to flow through column loaded amylose resin. Recombinant protein accumulated amylose resin was washed with column buffer (12 × volume). The rOfCfD-MBP fusion protein was eluted by applying elution buffer (Column buffer + 10 mM maltose). Simultaneously, MBP was overexpressed and purified as described above. Protein concentration was monitored by the Bradford assay (Bradford, 1976) and samples obtained at different purification steps were subjected to 12% SDS-PAGE along with molecular standards (Enzynomics, Seoul, Korea). The 0.05% Coomassie

blue R-250 was used to stain the gel, followed by gel was subjected to a standard de-staining procedure.

2.12. Protease assay

In order to measure the protease activity of rOfCfD-MBP, Pierce® Protease Assay Kit (Thermo Scientific, U.S.A) was used. Briefly, 100µL succinylated casein solution was added to one set of microplate wells as a substrate. 100µL assay buffer was added to a duplicate set of wells to serve as blanks. Then, 50µL of rOfCfD-MBP, MBP (negative control) and trypsin (positive control) protein samples were treated to both succinylated casein wells and corresponding blank wells. Plate was incubated for 20 minutes at room temperature (RT). After that, 50µL TNBSA working solution was added to each well and incubated for 20 minutes at RT. Measure absorbances of wells in a plate reader set to 450 nm. For each well calculate the change in absorbance at 450 nm by subtracting of the blank from that of the corresponding casein well. This absorbance difference was generated by the proteolytic activity of the protease.

2.13. Statistical analysis

All the data are presented as relative mRNA expression means ± standard deviation (SD). To determine statistical significance, the obtained data were subjected to statistical analysis using unpaired, two-tailed *t*-test to calculate the *P*-value using GraphPad program (GraphPad Software, Inc.). The significant difference was defined at $P < 0.05$.

3.0. Chapter I

Molecular characterization of complement component 1r (C1r) and complement component 1s (C1s) from classical pathway, as the member of initiator complex of complement system

Characterization of rock bream (*Oplegnathus fasciatus*) complement components, C1r and C1s, in terms of molecular aspects, genomic modulation and immune responsive transcriptional profiles upon bacterial pathogen

Abstract

C1r and C1s are complement components that plays crucial role via activation of complement classical cascade system. As classical initiator complex molecules in pathogen-induced signaling cascade, C1r and C1s modulate innate immunity. In order to understand the immune responses of teleost C1r and C1s, *Oplegnathus fasciatus* C1r and C1s genes (OfC1r and OfC1s) were identified and characterized. OfC1r genomic sequence enclosed with thirteen exons which represented 704 aa (78 kDa) putative peptide where, eleven exons of OfC1s represented 691 aa polypeptide (76 kDa). In addition, genomic analysis revealed that both OfC1r and OfC1s were located in a single chromosome. These putative polypeptides were demonstrated and composed of two CUB domain, EGF domain, two CCP domains and catalytically active serine protease domain. Phylogeny of C1r and C1s was studied and OfC1r evolutionary cladded to that of *Pundamilia nyererei* (identity = 73.4 %) where OfC1s to that of *Oryzias latipes* (identity = 58.2 %). Based on real-time qPCR analysis results, OfC1r and OfC1s transcript were detected along the eleven different tissues with higher levels of OfC1r in blood and OfC1s in liver. Putative role of OfC1r and OfC1s in response to pathogenic bacteria (*Edwardsiella tarda* and *Streptococcus iniae*) were profiled in liver tissues and OfC1r and OfC1s revealed up-regulated transcript levels. Overall findings of present study associate that OfC1r and OfC1s is involved in classical signaling cascade, but also contributes to the antibacterial responses.

Keywords: Synteny; genomic arrangement; classical system; immune responses

3.1. Introduction

The complement system is an essential part of the innate and adaptive immune system which plays crucial roles in eradicating of invaded pathogenic microorganisms by forming membrane attack complex (MAC) on their surfaces, in addition to the clearing of immune complexes, chemotaxis and inflammation (Nakao et al., 2011). Three semi-independent biochemical pathways such as classical, alternative and lectin pathway activate the complement system. The classical pathway is initiated by the activation of C1 multimolecular complex which is assembled with one subunit of recognition C1q consisting six heterotrimeric globular heads, two molecules each of C1r and C1s (Wang and Secombes, 2003). C1q is the key pattern recognition molecule that has been shown to bind with various activator structures like immune complexes, lipopolysaccharides, C-reactive proteins and apoptotic cells etc., in order to trigger the classical pathway (Gal et al., 2009). Due to such antibody-antigen interactions, the C1 complex changes the structural conformation that causes the autolytic activation of C1r. Subsequently, activate the C1s precursors and mediate initial proteolytic activation of C2 and C4 by releasing the small activation molecule (C2b and C4b) and major molecule (C2a and C4a) where C4b and C2a associated to form C3 convertase (C4bC2a) (Reid et al., 1998) (Holland and Lambris, 2002). Resulted C3 convertase in complement cascade system ultimately release of anaphylatoxins (C3a, C5a) and continue formation of MAC molecules on surface of the microbes thus mediate phagocytosis and cytolysis of pathogen (Carroll and Georgiou, 2013).

The early complement protease C1 complex members of C1r/C1s in the classical pathway belongs to the serine proteases (Matsushita and Fujita, 1992) which have been under investigated from teleost. Structurally, C1r and C1s comprised of

five characteristics domains in two functionally distinct fragments. The N-terminal interaction fragment of C1r and C1s (CUB1-EGF-CUB2) mediates calcium dependent C1q-$C1r_2C1s_2$ affiliation (Busby and Ingham, 1990) (Kardos et al., 2008). Important C-terminal catalytic region, which is CCP1-CCP2-SPr (CCP; complement control protein, SPr; serine protease) make bond between C1r and C1s via disulfide bond to establish the enzymatically active properties (Arlaud and Gagnon, 1985). The catalytic fragment of C1r molecule involve in formation of homodimer by facilitating the core for C1s-C1r-C1r-C1s tetramer (Arlaud et al., 2002b) (Lacroix et al., 2001). Further CCP2 domain probably enhances the catalytic activity by binding to the SPr domain (Arlaud et al., 2002b) (Gaboriaud et al., 2000). An immune related essential feature of C1r and C1s which possessed complement system activation associated with C1q peptide. Upon classical pathway activation, it creates an important adaptive immune response (Wang and Secombes, 2003). However, activation and activity of C1r and C1s molecules are assured by the serpin C1-inhibitor towards the balance of the complement activity and thereby maintain the tissue homeostasis (Beinrohr et al., 2007) (He et al., 1997) in the absence of pathological situations. Interestingly, the genes encoding human C1r and C1s have been showed a tail-to-tail arrangement in a single chromosome, suggesting that they were derived from tandem duplication of a common ancestor gene (Kusumoto et al., 1988) in order to possess their protease activity together.

The complete sequence of C1r and C1s has been documented and characterized only from human (Journet and Tosi, 1986) (Mackinnon et al., 1987), and murine (Circolo et al., 2003) (Sakai et al., 1998). Up to date, molecular characterization of teleostean C1r/C1s is much limited, though sequence information of common carp C1r/C1s (Nakao et al., 2001) and sequencing and expression of

rainbow trout C1r (Wang and Secombes, 2003) have been reported. In fact teleost C1r and C1s genes characterization in molecular and functional basis, remain to be documented. The rock bream (*Oplegnathus fasciatus*) is an economically crucial aquacrop in Korea, because of its high market price (Lipton, 2010). However, rock bream production losses have been reported from aquaculture industry in Korea due to growing trend to its vulnerability to infection caused by bacterial pathogens (Park Ii, 2009). Therefore, in an effort to identify the molecular defense mechanism of immune genes involved in antibacterial defenses and the gene expression profile of bacteria challenged rock bream could be a productive approaches during disease prevention in rock bream industry. Hence, rock bream C1r (OfC1r) and C1s (OfC1s) have remarkably overlooked towards elucidating its specific biosynthetic aspects.

This study we support the knowledge to understand the molecular insights of OfC1r and OfC1s into the immune mechanism involved in rock bream by means of: (1) characteristic properties at molecular level, (2) cloning the rock bream C1r and rock bream C1s at genomic DNA level and depict the genomic arrangements and (3) spatial mRNA profiles in different healthy tissues and temporal mRNA expression upon immune stimulation.

3.2. Results

3.2.1. Sequence characterization and domain architecture analysis

The putative nucleotide sequences of OfC1r and OfC1s were identified through BLASTX homology screening and indicated that they are homologous counterparts of vertebrate C1r and C1s. Subsequently, the OfC1r and OfC1s sequence information have been deposited in GenBank under accession numbers of KP241694 and KP241695, respectively. The OfC1r complete cDNA sequence was 2533 bp in

length, composed of a putative CDS of 2115 bp. The 5' untranslated region (UTR) and 3' UTR lengths were 105 bp and 313 bp, respectively. The CDS corresponding to OfC1r could be coded a putative polypeptide of 704 aa with its predicted molecular mass and theoretical isoelectric point (*p*I) were 78 kDa and 5.3, respectively (Fig. 3.1). Whereas, the OfC1s full length cDNA (2724 bp) sequence bears 5' UTR (32 bp), 3' UTR (616 bp) and CDS with 2076 bp which translated into 691 aa of putative polypeptide. Its predicted molecular mass was 76 kDa and theoretical *p*I was 5.2 (Fig. 3.2). Interestingly, deduced peptide sequence of OfC1r and OfC1s were comprised similar domain architecture including a signal peptide, two CUB domains, calcium binding EGF domain, two complement control protein (CCP) domains and serine protease trypsin domain (SPr) (Fig 3.3). Apart from that, cleavage site and few N-linked glycosylation sites were identified.

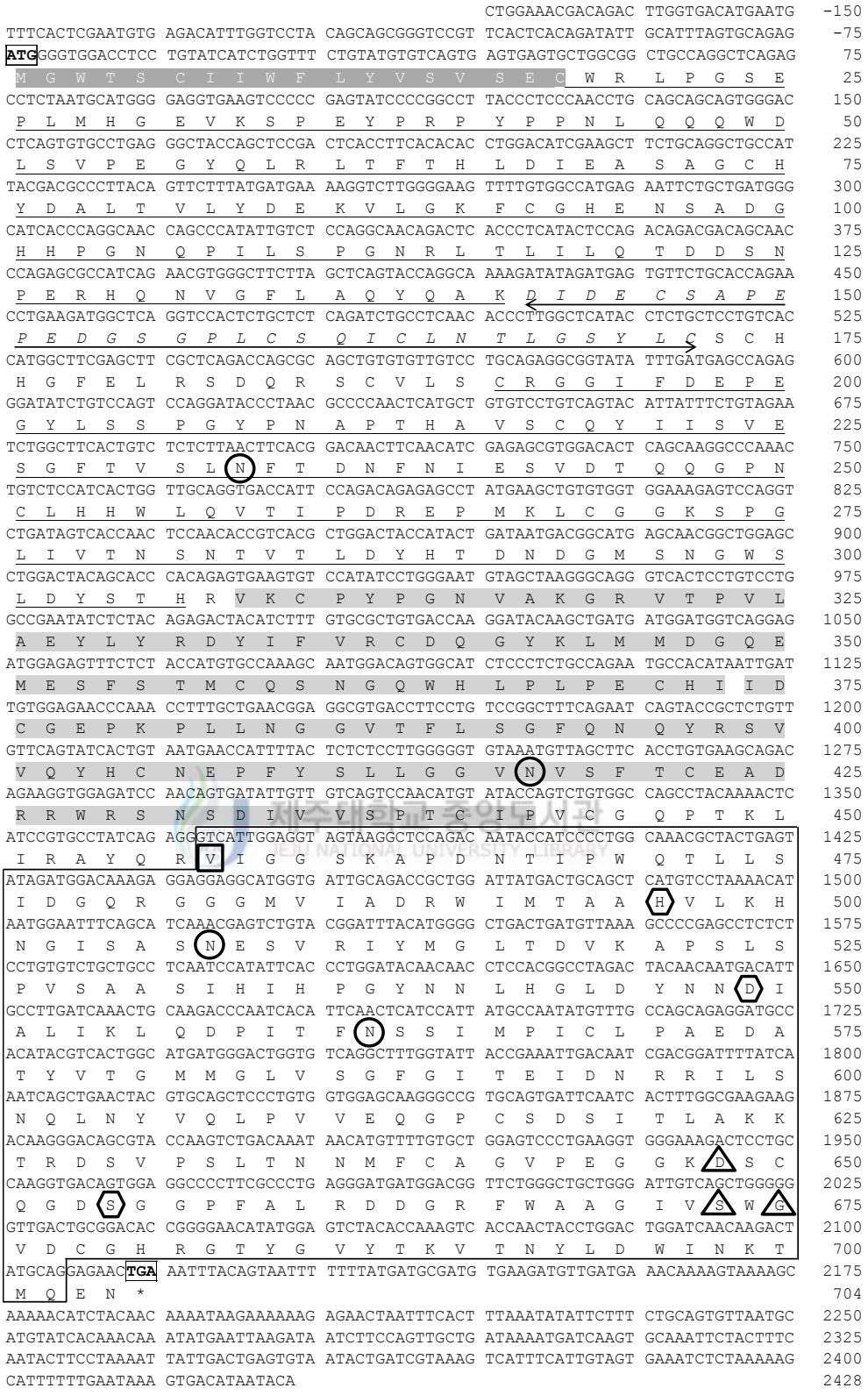


Figure 3.1 Nucleotide and deduced amino acid sequences of OfC1r. The nucleotide sequence (upper) and the deduced amino acid sequence (lower) are numbered. The start codon (ATG) and stop codon (TGA) are boxed. The SP is dark grey shaded. CUB domains are underlined, CCP domains are light grey shaded, Ca binding EGF domain is italicized and marked in arrow, Trypsin domain is boxed, cleavage site is mark in square, active sites are mark in hexagonal, substrate binding sites are mark in triangle and N-linked glycosylation sites are mark in circle.

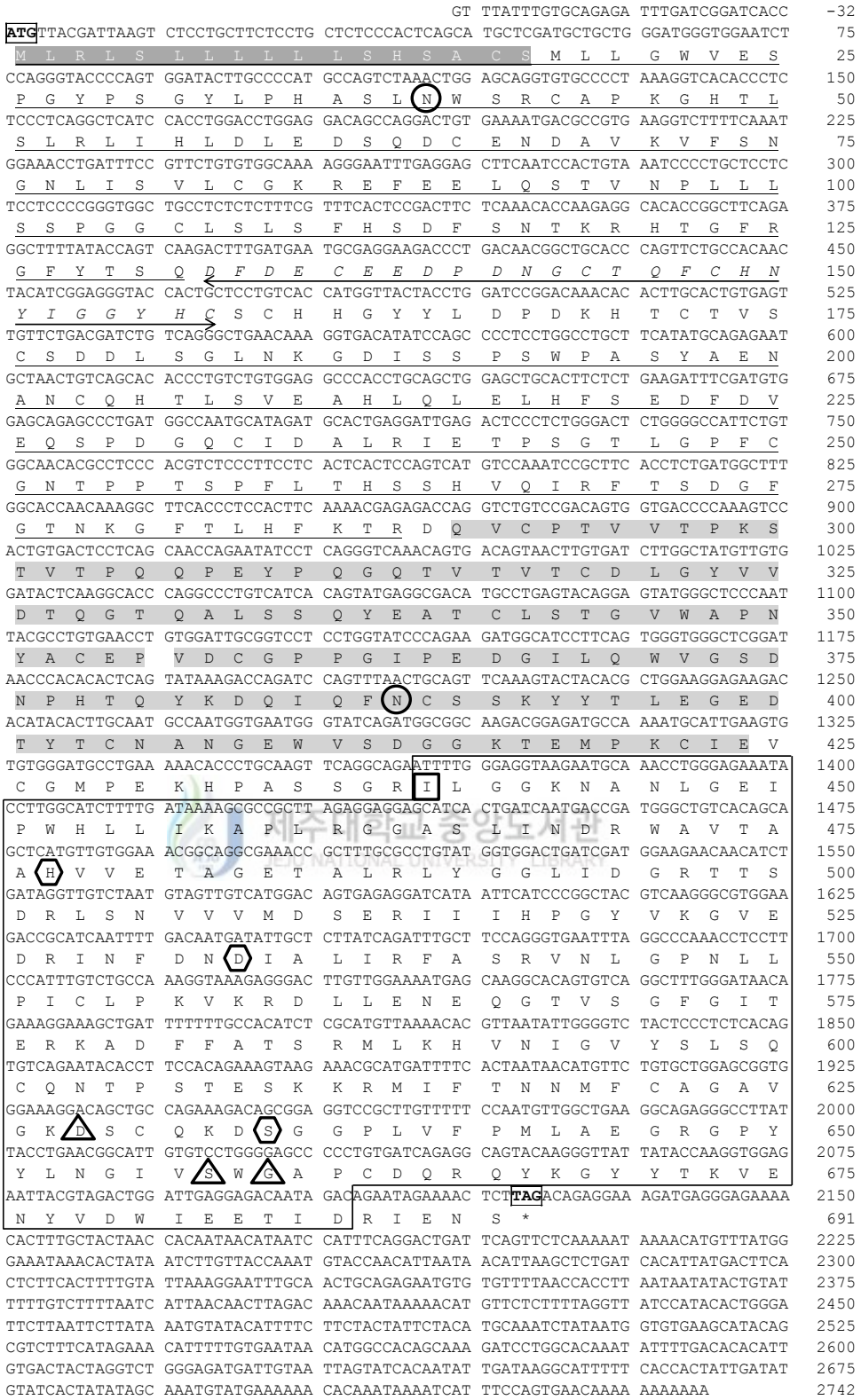


Figure 3.2 Nucleotide and deduced amino acid sequences of OfC1s. The nucleotide sequence (upper) and the deduced amino acid sequence (lower) are numbered. The start codon (ATG) and stop codon (TGA) are boxed. The SP is dark grey shaded. CUB domains are underlined, CCP domains are light grey shaded, Ca binding EGF domain is italicized and marked in arrow, Trypsin domain is boxed, cleavage site is mark in square, active sites are mark in hexagonal, substrate binding sites are mark in triangle and N-linked glycosylation sites are mark in circle.

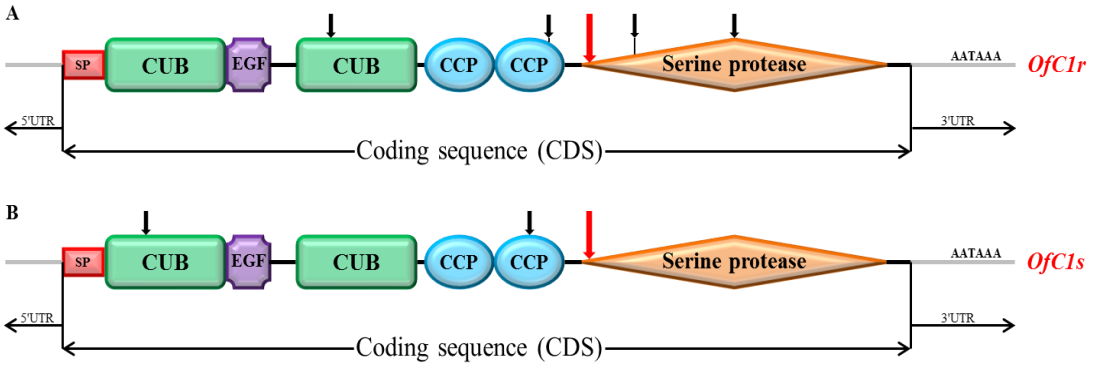
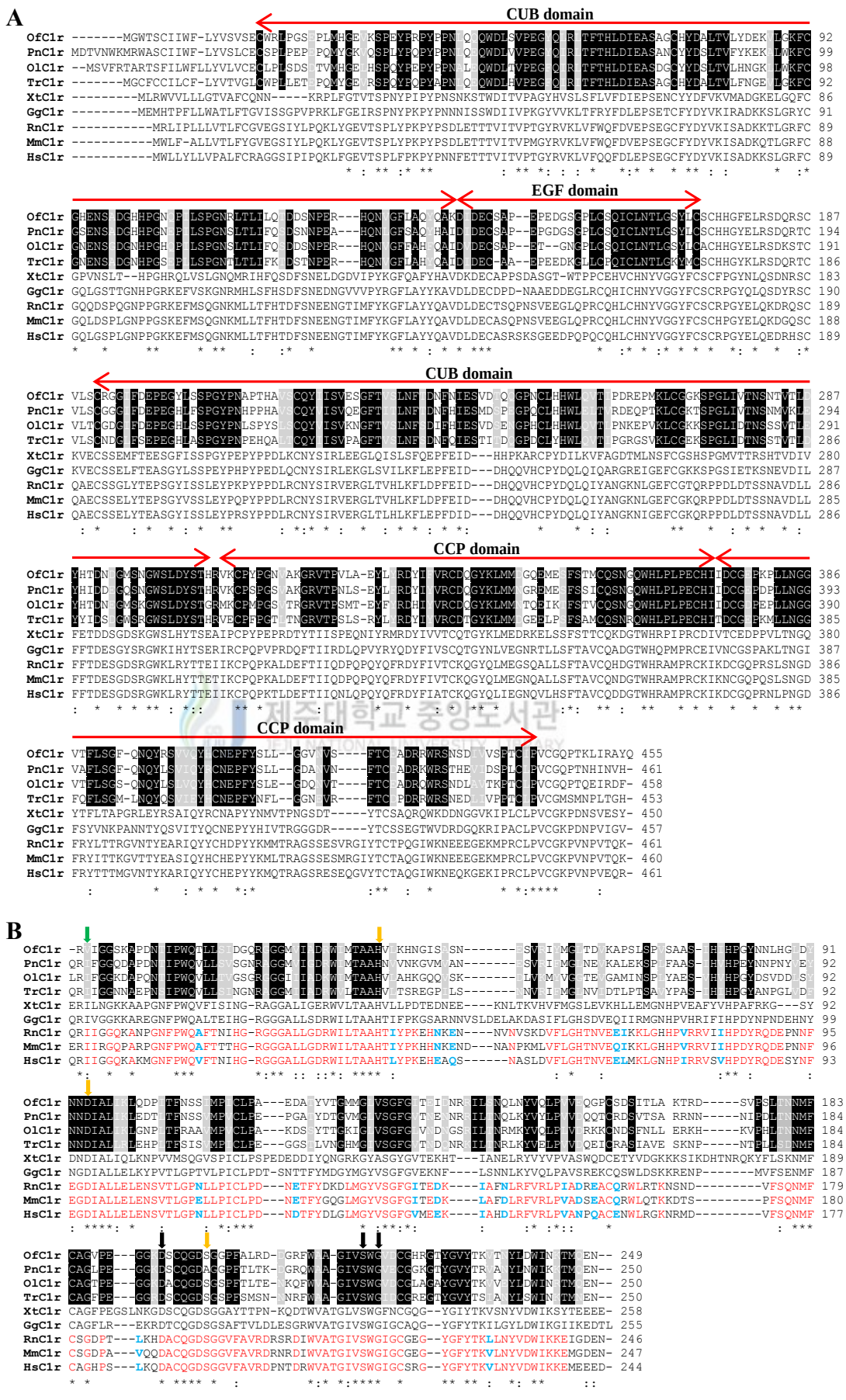
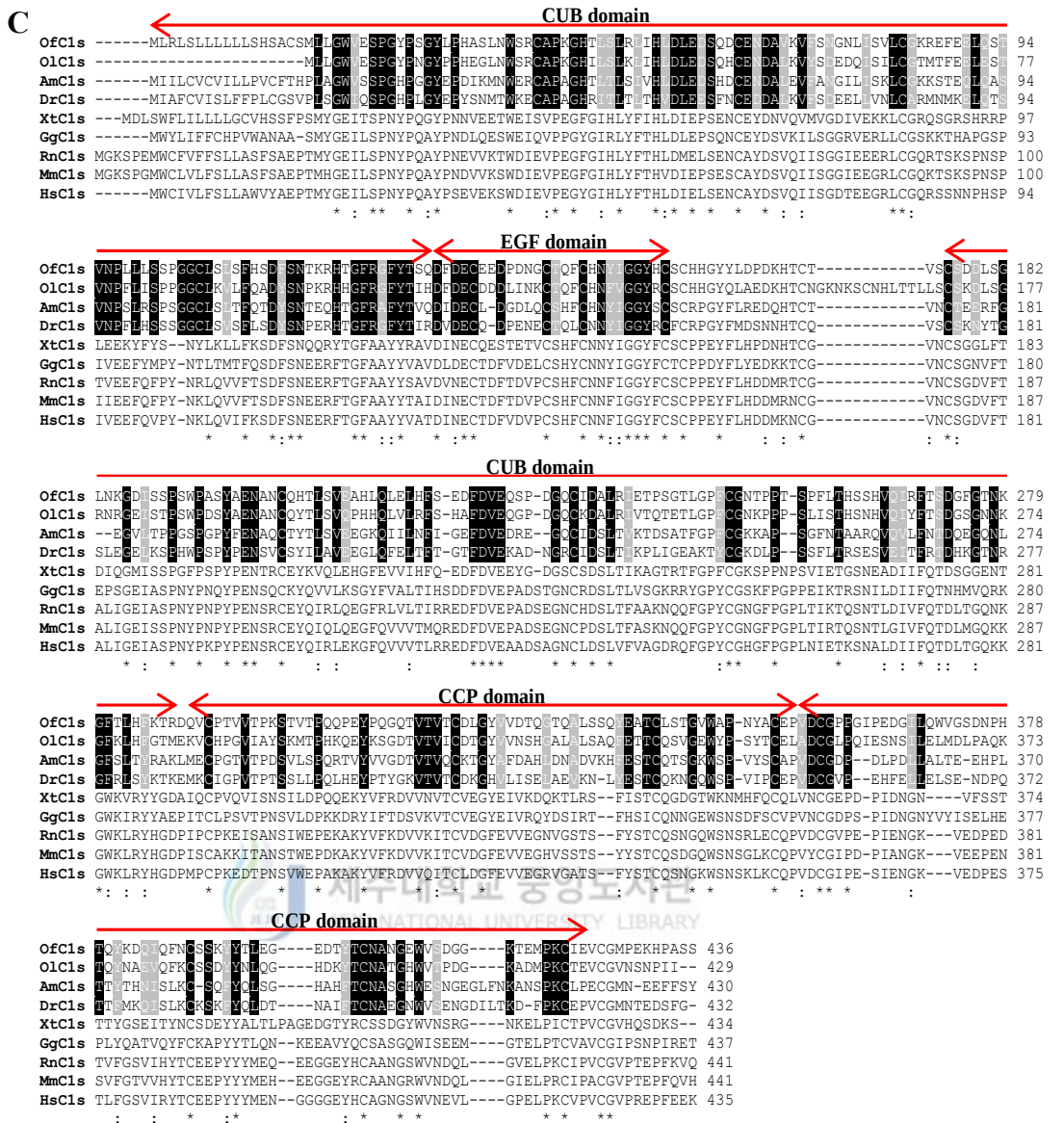


Figure 3.3. The domain architecture of rock bream C1r (A) and C1s (B). Conserved domains denoted by different symbols within the coding region. Black dashes represent the peptide sequences. Red arrow corresponds to cleavage site on activation. Black arrows indicate the N-linked glycosylation sites. SP; signal peptide, CUB; CUB domain, EGF; calcium binding EGF like domain, CCP; complement control protein domain and Serine protease trypsin domain.

3.2.2. Homology amino acid sequence analysis

BLASTP homology sequence analysis of OfC1r and OfC1s showed that they were aligned with other known C1r and C1s homologous sequences deposited in NCBI sequence database. These results were further affirmed by multiple sequence alignment of nine homology sequence from different species (Fig. 3.4). The multiple sequence alignment analysis revealed more conserved residues within the functionally active serine protease domain corresponding to the OfC1r (Fig. 3.4B) and OfC1s (Fig. 3.4D). In addition, three active sites and three substrate binding sites were detected in SPr domain where substrate binding sites were identical among selected C1r and C1s species. The two active sites were conserved among C1r sequences while all three active sites were conserved among C1s sequences. Moreover, multiple sequence alignment of CUB1-EGF-CUB2-CCP1-CCP2 domain region showed that sequence homology within the fish counterparts (Fig. 3.4A; C1r and 3.4B; C1s).




C
CUB domain
EGF domain
CUB domain
CCP domain
CCP domain
OfC1s
OlC1s
AmC1s
DrC1s
XtC1s
GgC1s
RnC1s
MmC1s
HsC1s

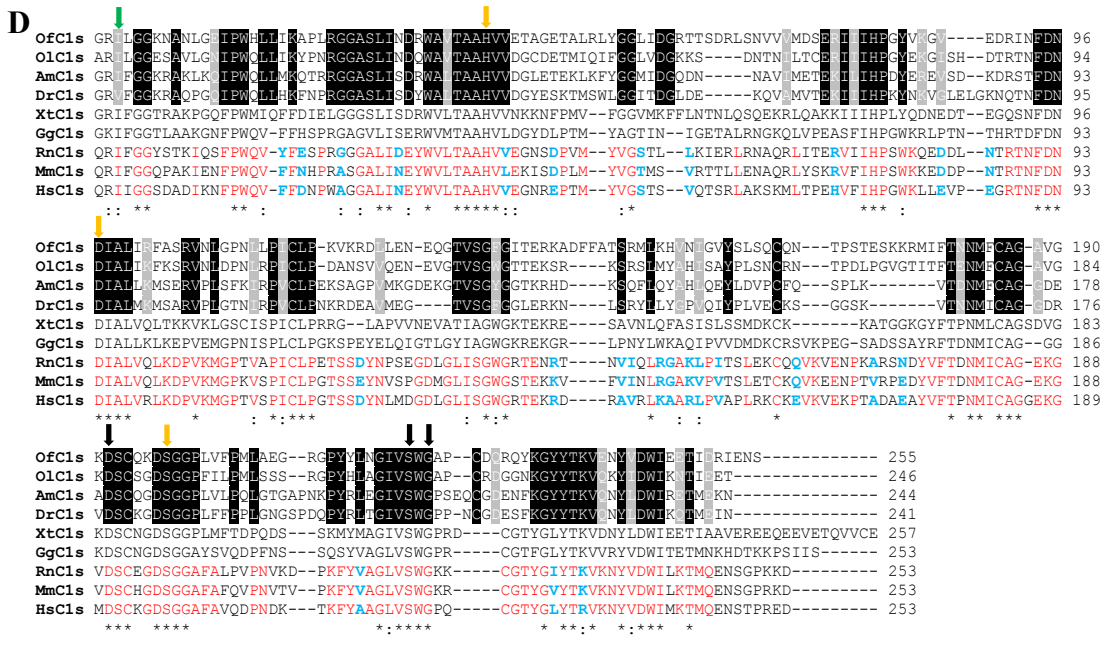

D
OfC1s
OlC1s
AmC1s
DrC1s
XtC1s
GgC1s
RnC1s
MmC1s
HsC1s

Figure 3.4. Multiple sequence alignment of OfC1r (A, B) and OfC1s (C, D) with other known C1r and C1s amino acid sequences using ClustalW (v2.1). Identical residues are denoted by (*) and similar residues are denoted by (:). Identical residues within fish group shaded by black color and similar residues are shaded by grey color. Missing amino acids are indicated by dashes. The C1r and C1s corresponding CUB1-EGF-CUB2-CCP1-CCP2 domain region multiple alignment represent by A and C, respectively. The C1r and C1s corresponding serine protease domain multiple alignment represent by B and D, respectively. Cleavage site is marked by green color arrow. Active sites are marked by yellow color arrow and black color arrow represented the substrate binding sites.

The pairwise comparison of amino acid sequences of OfC1r and other orthologs supported the conservation of the C1r in different lineages (Table 3.1). The overall sequence identity of OfC1r ranged from 73.4 % to 39 % where it shared highest identity (73.4 %) and similarity (85.4 %) to the *Pundamilia nyererei* C1r ortholog. In addition, it also demonstrated quite low identity with higher vertebrate C1r orthologs.

Table 3.1: Homology analysis of identity and similarity percentage from *OfC1r*

	C1r	Identity %																	
		O. fasciatus	*P. nyererei*	*O. latipes*	*T. rubripes*	*X. maculatus*	*L. oculatus*	*C. carpio*	*L. chalumnae*	*X. tropicalis*	*P. bivittatus*	*G. gallus*	*A. carolinensis*	*O. orca*	*R. norvegicus*	*M. musculus*	*M. gallopavo*	*H. sapiens*	*B. taurus*
Similarity %	*Oplegnathus fasciatus*		**73.4**	**68.2**	**67.4**	**67.4**	**52.3**	**49.0**	**42.5**	**41.0**	**40.4**	**40.1**	**39.9**	**39.9**	**39.9**	**39.9**	**39.5**	**39.3**	**39.0**
	Pundamilia nyererei	**85.4**		66.4	64.7	64.5	53.1	47.5	40.9	39.8	40.2	38.0	40.7	36.9	37.3	37.0	37.8	37.2	36.3
	Oryzias latipes	**84.2**	82.1		60.5	63.7	49.9	47.9	40.8	39.2	38.8	39.5	39.3	36.1	36.5	37.1	38.4	36.4	35.5
	Takifugu rubripes	**81.5**	79.7	77.7		58.4	48.5	47.2	41.4	39.5	39.9	40.2	39.4	38.2	40.4	39.3	39.9	38.3	39.6
	Xiphophoru maculatus	**82.8**	80.6	81.5	75.1		49.0	45.8	39.0	38.2	39.6	39.0	38.7	36.3	36.5	36.8	38.4	36.0	36.7
	Lepisosteus oculatus	**71.0**	70.3	68.9	69.6	67.2		47.1	42.5	43.1	44.1	42.7	43.5	42.1	42.7	41.8	42.1	43.0	42.6
	Cyprinus carpio	**66.3**	66.1	65.0	64.3	64.1	64.2		38.8	36.6	38.2	37.1	36.9	36.5	36.8	37.1	37.1	36.2	36.8
	Latimeria chalumnae	**61.6**	61.3	62.0	60.6	57.6	63.8	58.7		45.5	47.4	49.0	46.8	46.6	46.6	46.4	48.1	46.3	46.9
	Xenopus tropicalis	**60.7**	59.8	60.6	58.8	58.2	62.7	56.4	63.6		50.6	51.0	51.5	49.5	49.0	48.6	50.8	51.2	48.8
	Python bivittatus	**60.0**	61.0	61.1	59.3	59.2	62.3	56.7	66.0	68.4		64.2	77.1	58.3	56.9	57.0	64.2	58.1	56.3
	Gallus gallus	**59.8**	60.3	61.0	61.5	58.1	64.6	56.2	65.9	68.7	79.2		65.8	61.8	60.4	59.9	93.4	60.5	60.7
	Anolis carolinensis	**59.8**	60.5	59.9	59.1	56.5	62.2	56.3	66.6	70.7	88.6	80.3		58.2	58.7	57.9	65.5	60.0	57.1
	Orcinus orca	**60.4**	57.9	58.6	59.9	57.0	62.8	56.6	66.0	68.4	75.2	76.1	74.3		82.9	80.5	61.6	85.5	84.7
	Rattus norvegicus	**60.4**	59.9	57.8	61.4	57.1	63.2	56.7	66.8	67.5	73.7	76.7	75.5	89.4		90.5	59.7	83.2	78.2
	Mus musculus	**59.8**	59.2	58.1	60.7	56.9	62.9	56.6	65.8	67.9	74.8	76.1	74.4	88.3	94.3		59.1	80.9	76.7
	Meleagris gallopavo	**59.4**	59.3	60.7	61.0	57.9	63.9	56.2	65.7	69.0	79.1	97.1	80.4	76.7	76.4	75.4		60.9	59.9
	Homo sapiens	**59.6**	59.2	59.3	59.4	57.6	63.1	56.5	66.0	68.6	74.0	77.1	74.8	92.8	90.7	89.4	77.1		81.1
	Bos taurus	**59.7**	57.8	57.5	60.0	57.0	62.0	56.5	65.4	67.1	73.0	75.8	74.0	92.6	87.7	87.0	75.8	90.4	

Based on pairwise homology analysis, the highest homology of OfC1s was demonstrated highest identity (58.2 %) and similarity (70.8 %) to the *Oryzias latipes*

C1s ortholog (Table 3.2). Moreover, the overall sequence identity of OfC1s ranged from 58.2 % to 33.6 % among them at amino acid and nucleotide levels.

Table 3.2: Homology analysis of identity and similarity percentage from *OfC1s*

C1s		Identity %																	
		O. fasciatus	*O. latipes*	*A. mexicanus*	*D. rerio*	*X. tropicalis*	*X. laevis*	*A. carolinensis*	*L. chalumnae*	*S. acanthias*	*R. norvegicus*	*A. sinensis*	*H. sapiens*	*M. musculus*	*M. fascicularis*	*S. scrofa*	*B. taurus*	*G. gallus*	*M. gallopavo*
Similarity %	*Oplegnathus fasciatus*		**58.2**	**47.4**	**46.7**	**37.6**	**36.4**	**36.4**	**36.1**	**35.8**	**35.7**	**35.1**	**35.1**	**35.0**	**34.9**	**34.8**	**34.7**	**33.6**	**33.6**
	Oryzias latipes	**70.8**		45.7	44.9	36.7	35.3	38.2	35.8	33.9	35.2	36.2	34.8	34.1	35.5	35.5	35.7	35.2	35.3
	Astyanax mexicanus	**65.1**	64.7		54.7	38.2	38.6	37.3	37.9	37.2	35.5	37.9	35.2	34.3	35.3	36.3	35.1	36.4	36.2
	Danio rerio	**64.3**	62.4	70.9		36.6	36.2	35.9	36.5	35.7	37.0	36.5	35.3	34.3	35.6	35.1	34.6	36.6	35.9
	Xenopus tropicalis	**56.1**	55.6	57.7	57.5		80.0	48.6	45.7	44.6	48.7	49.6	49.6	48.6	49.8	47.8	47.0	47.2	47.2
	Xenopus laevis	**55.6**	54.3	57.5	57.2	88.9		48.2	46.9	43.7	48.2	48.2	49.7	48.1	49.4	47.7	47.1	45.6	45.7
	Anolis carolinensis	**54.1**	56.1	55.5	54.2	66.0	67.1		45.8	42.5	55.4	63.1	57.4	53.5	57.4	56.7	54.3	61.6	61.1
	Latimeria chalumnae	**55.2**	55.6	56.5	55.4	64.4	64.6	65.2		44.7	47.0	47.8	48.9	47.8	48.2	47.7	45.7	46.6	46.3
	Squalus acanthias	**57.5**	55.3	55.8	55.9	65.4	65.4	61.6	65.9		42.7	42.5	42.0	42.4	42.0	40.6	40.6	40.6	40.4
	Rattus norvegicus	**55.5**	54.5	55.0	55.2	67.7	67.4	72.0	64.6	61.7		57.1	75.4	82.6	76.0	75.5	71.8	52.6	52.3
	Alligator sinensis	**56.1**	55.7	56.3	56.4	67.5	67.7	77.1	66.4	62.0	74.2		58.5	56.6	58.9	58.6	56.0	66.9	65.5
	Homo sapiens	**54.1**	53.3	55.2	56.1	68.2	68.5	73.3	65.6	61.9	85.9	74.6		73.4	94.2	78.8	74.2	55.3	54.4
	Mus musculus	**54.9**	53.2	54.9	53.0	68.2	66.9	70.0	65.0	61.1	91.9	72.9	84.3		72.7	72.6	69.8	52.9	53.1
	Macaca fascicularis	**54.1**	54.2	56.3	57.1	68.9	69.2	73.1	65.0	61.2	86.2	74.7	96.7	83.9		78.4	75.1	55.5	54.5
	Sus scrofa	**53.4**	55.3	55.2	55.5	66.3	66.8	71.6	64.9	60.6	86.0	73.6	87.2	84.7	87.9		82.0	54.3	54.0
	Bos taurus	**53.0**	53.8	54.0	54.5	65.6	66.0	70.5	64.7	59.4	84.3	74.5	85.9	82.9	86.6	90.1		52.8	52.8
	Gallus gallus	**54.4**	56.7	55.9	55.8	67.1	66.4	75.5	66.2	62.6	71.9	80.7	73.2	72.6	72.6	73.3	72.4		88.4
	Meleagris gallopavo	**54.9**	55.0	55.6	54.7	65.9	66.9	75.3	65.4	61.9	72.1	79.9	72.4	72.3	72.6	72.4	73.0	92.9	

Molecular phylogenetic tree was constructed to delineate the evolutionary relationship among C1r and C1s counterparts (Fig. 3.5). The tree have cladded into two main branches: one main branch was assembled with C1r homologous and the other was grouped with C1s homologous members. Each cluster was included several sub families admitting fish, amphibian, reptile, aves and mammals of corresponding counterparts. Within C1r and C1s main clusters, fish C1r and C1s composed a sub-clade (except *Latimeria chalumnae* and *Squalus acanthias*), respectively and the other entire vertebrate create a second sub-clade. OfC1r was positioned to that of *Pundamilia nyererei* and where OfC1s was placed to that of *Oryzias latipes* with a strong bootstrap replication of 93 and 100, respectively.

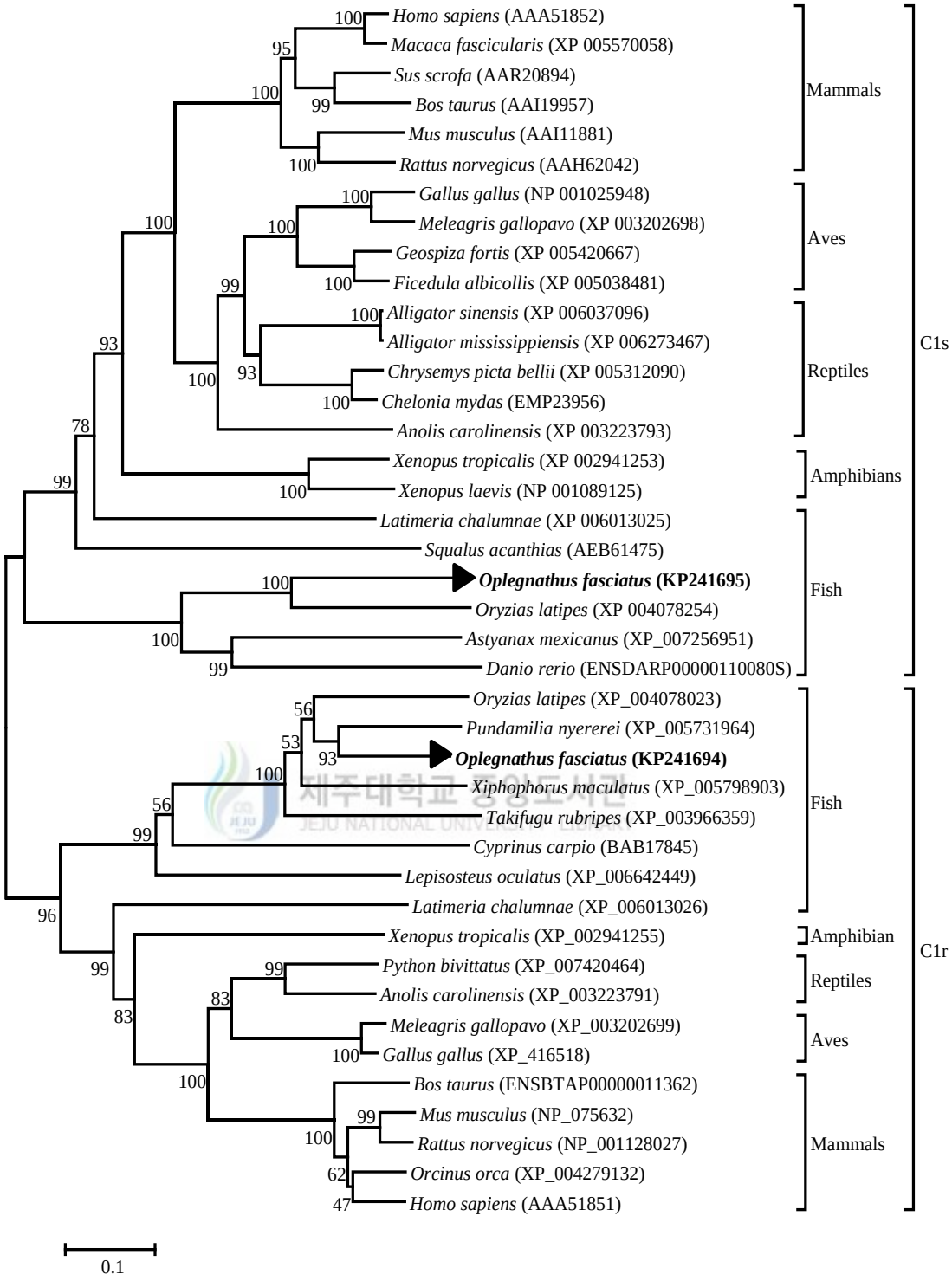
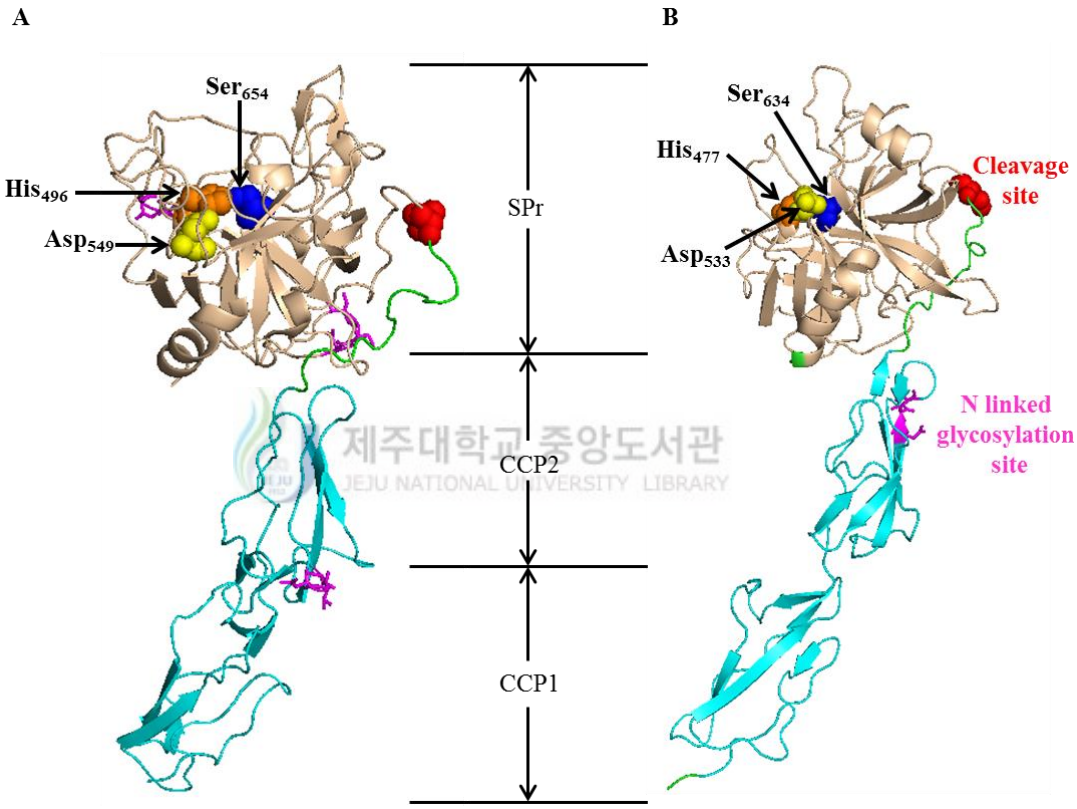


Figure. 3.5. Phylogenetic analysis of OfC1r and OfC1s with their known orthologues. The tree is based on the alignment of the full-length amino acid sequences. The tree was constructed using ClustalW (v2.1) and MEGA (v5.0). The numbers at the branches denote the bootstrap replications for 5000 replicates. The GenBank accession numbers are shown within brackets next to each species.

The partial monomeric tertiary structure (CCP1-CCP2-SPr region) of OfC1r (Fig. 3.6A) and OfC1s (Fig. 3.6B) was graphically illustrated based on the X-ray structure of its corresponding human C1r (PDB, 1gpz; 308-702 aa) and C1s (PDB, 1zjk; 288-688 aa) crystal structures. Apart from that, OfC1r create homodimer (Fig. 3.6C) while making a core for C1s-C1r-C1r-C1s tetramer formation. Interestingly, three active sites were organized to each other in the SPr domain.



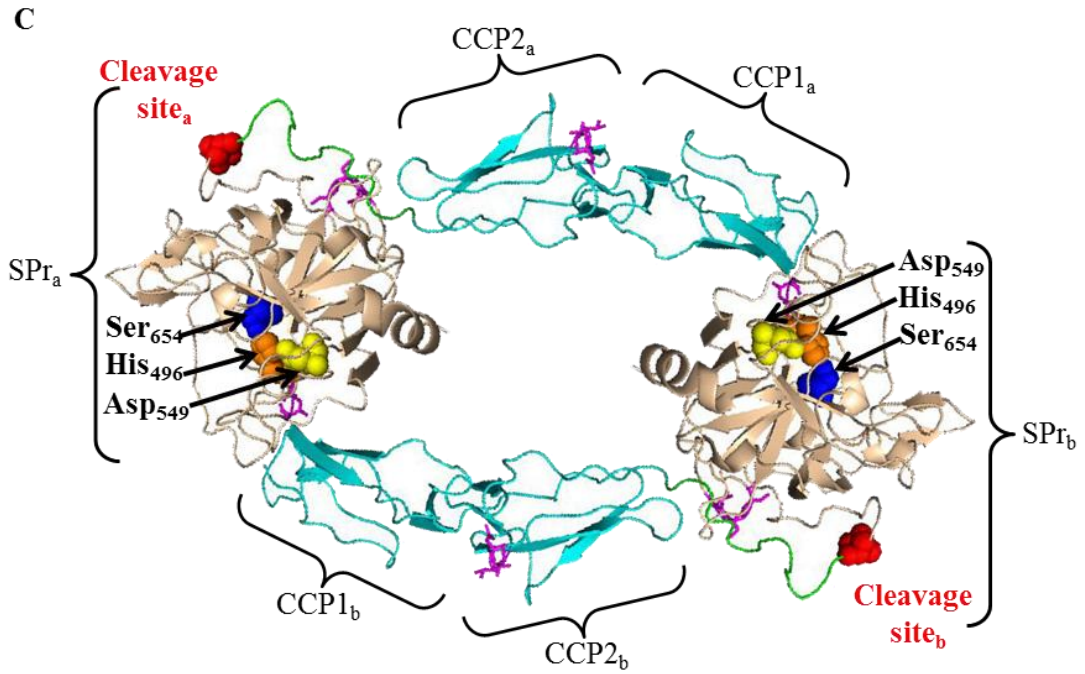


Figure. 3.6. Comparative monomeric tertiary structure of OfC1r (A) and OfC1s (B). The conserved active sites are shown in yellow color spheres. Cleavage site is marked by red color spheres. N-linked glycosylation sites are colored in purple color. Serine protease domain is in brown color. CCP domains are in blue color. The dimeric structure of OfC1r (C). One OfC1r monomer in green color and another OfC1r monomer in blue color

3.2.3. Comparative genomic structural analysis of C1r and C1s homologues

Genomic DNA sequences of *OfC1r* (Accession No: xxxxx) and *OfC1s* (Accession No: xxxxx) were analyzed to determine the probable intron-exon boundaries (Fig. 3.7). The genomic structure of *OfC1r* revealed in which thirteen exons are interrupted by twelve introns (Fig. 3.7A) whereas, *OfC1s* genomic structure demonstrated that eleven exons are interrupted by ten introns (Fig. 3.7B). All the exon-intron boundaries were proven the donor-acceptor (GT-AG) splicing rule. The *C1r* members from fish showed thirteen exons whereas, other vertebrate genomic structures comprised with eleven exons. Moreover, ten exons within the CDS of *C1r* members were conserved among the fish species where seven exons in the CDS were conserved among the mammalian group.

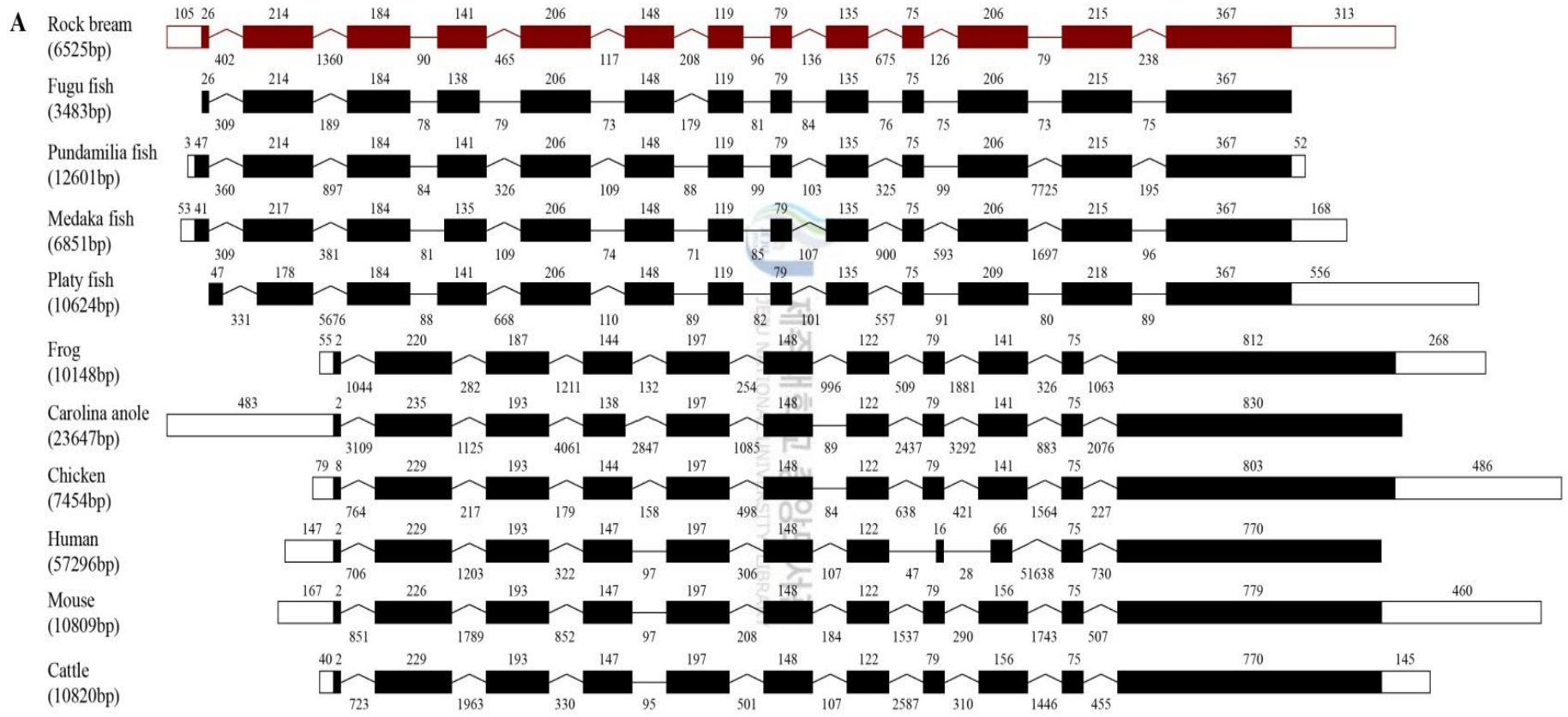


Figure 3.7A. Schematic genomic structures of C1r (rock bream: KP241694, fugu fish: NC_018896, pundamilia fish: NW_005187446, medaka fish: NC_019874, platy fish: NW_005372208, frog: NW_004668240, carolina anole: ENSACAT00000007508, chicken: NC_006088, human: ENST00000542285, mouse: ENSMUST00000068593, cattle: ENSBTAT00000011362) from various vertebrates. Coding regions are shown as dark shaded boxes. The 5′- and 3′- untranslated regions are denoted with empty boxes, and introns are indicated by lines. The sizes of the exons are shown above the exon boxes, and the intron sizes are shown below the intron lines.

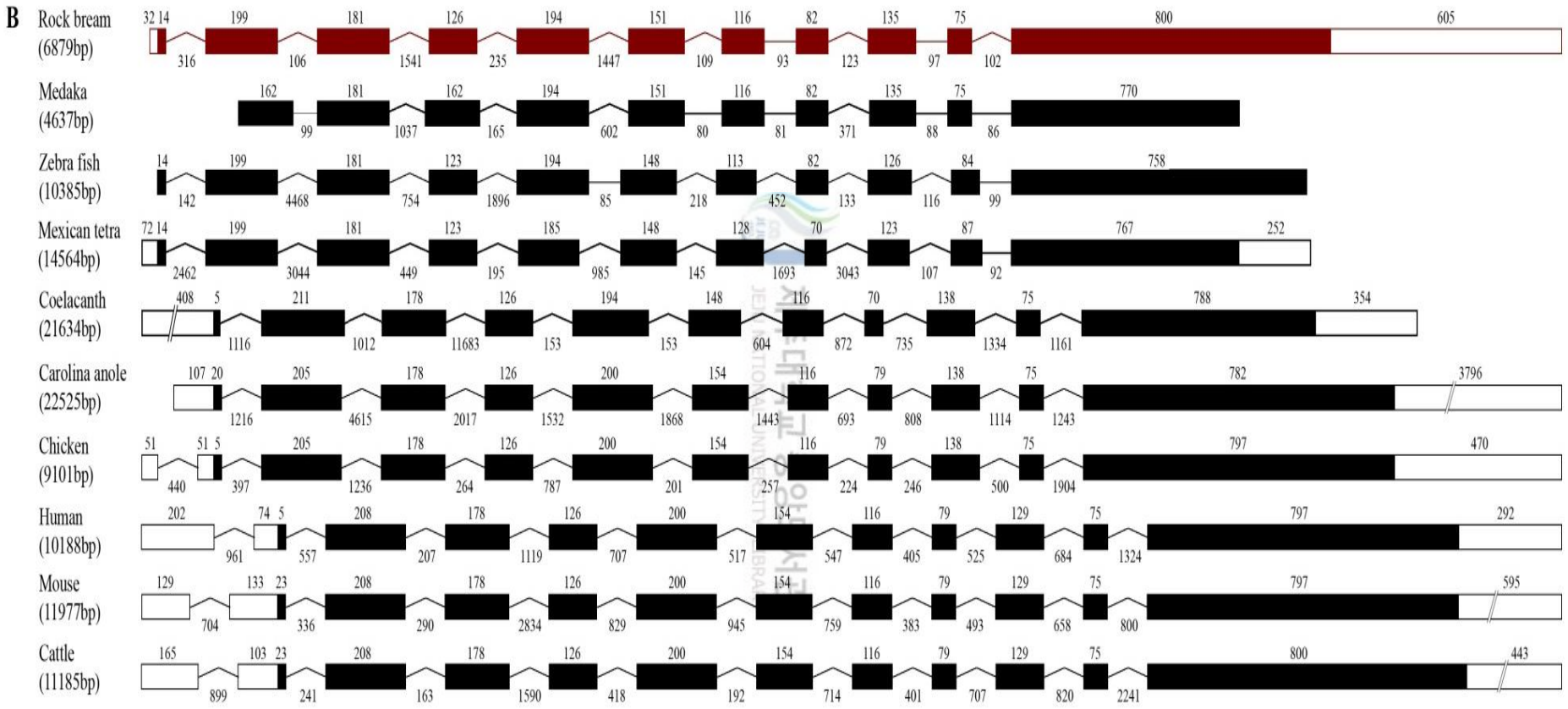


Figure 3.7B. Schematic genomic structures of C1s (rock bream: KP241695, medaka fish: NC_019874, zebrafish: ENSDART00000122049, mexican tetra: NW_006749413, coelacanth: ENSLACT00000026639, carolina anole: ENSACAT00000008080, chicken: ENSGALT00000023545, human: ENST00000360817, mouse: ENSMUST00000160505, cattle: ENSBTAT00000006358) from various vertebrates. Coding regions are shown as dark shaded boxes. The 5′- and 3′- untranslated regions are denoted with empty boxes, and introns are indicated by lines. The sizes of the exons are shown above the exon boxes, and the intron sizes are shown below the intron lines.

On the other hand, the genomic structures of *C1s* varied among different animal origins. Fish species (except medaka) consisted with eleven exons whereas; aves and mammals presented twelve exons. In contrast, first exon of the chicken and mammals represented the part of 5'UTR, in which second exon indicated the rest part of 5'UTR and part of CDS. Moreover, nine exons have shown similar in length within the mammalian *C1s* counterpart.

3.2.4. Synteny analysis

Synteny analysis revealed that rock bream consisted with one C1s and two C1r genes lie on a single chromosome in the same direction (Fig. 3.8). Note that, all investigated fish species have shown similar pattern of their C1r/C1s genes location on a same chromosome for same direction. However, higher vertebrate like aves, amphibians and mammals have depicted that opposite direction arrangement of C1r/C1s genes in one chromosome. Furthermore, different organisms were consisted with different numbers of C1r and C1s molecule combinations.

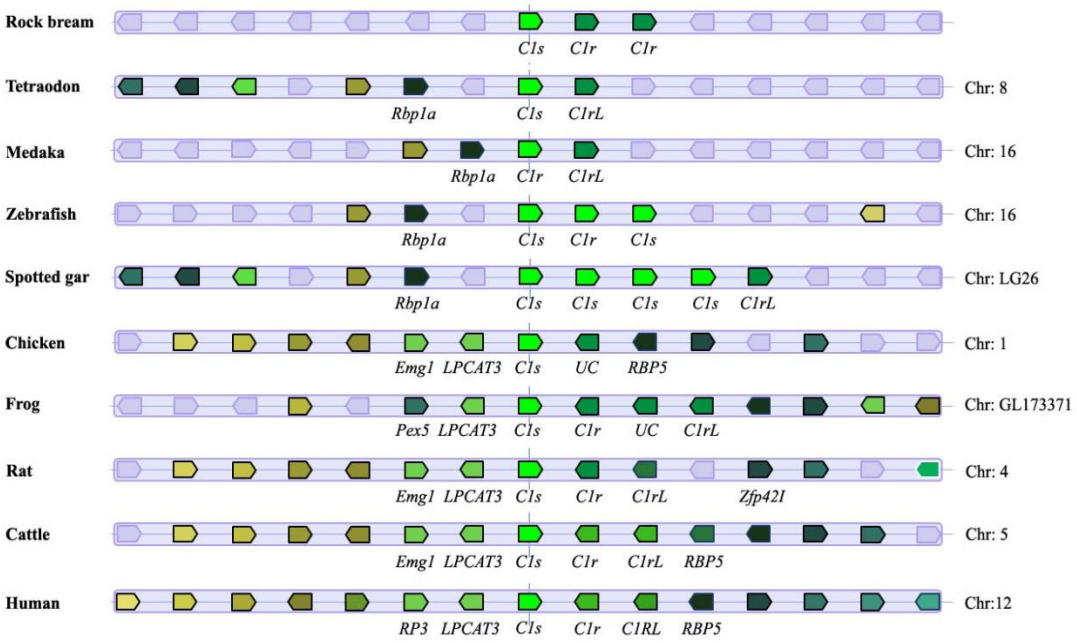


Figure 3.8. Physical synteny maps comparing the orthologues of C1r and C1s locus and gens flanking them in tetraodon, medaka, zebrafish, spotted gar, chicken, frog, rat, cattle and human. Identified genes (Rbp1a; retinol binding protein 1a, Emg1; EMG1 N1-specific pseudouridine methyltransferase, LPCAT3; Lysophosphatidylcholine acyltransferase 3, RBP5; retinol binding protein 5, Pex5; peroxisomal biogenesis factor 5, Zfp42l; Zinc finger protein 42-like, RP3; highly similar to Probable ribosome biogenesis protein NEP1, UC; un-characterized and L;like) are shown in different colors with the coding direction designated by the pointed end. Non-syntenic genes are indicated in purple color. The gene names are mentioned under the synteny map, species names are mentioned on the left side and chromosome location on right side. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

3.2.5. Constitutive transcriptional profiles in healthy rock breams

Spatial mRNA expression of OfC1r and OfC1s in healthy tissues revealed that constitutive transcriptional at detectable levels, with differential elevations in eleven different tissues (Fig. 3.9). Constitutive mRNA expression of OfC1r was significantly higher in blood (fold-1334.06), liver (fold-872.81), gill (fold-474.66) and kidney (fold-396.43) compare to muscle (Fig. 3.9A). Whereas, OfC1s was predominantly expressed in liver (fold-59.61) followed by, spleen (fold-42.88), kidney (fold-28.56) and head kidney (fold-22.76) compare to muscle (Fig. 3.9B). The other examined tissues were poorly transcribed where muscle tissue showed marginal mRNA expression for both OfC1r and OfC1s transcripts.

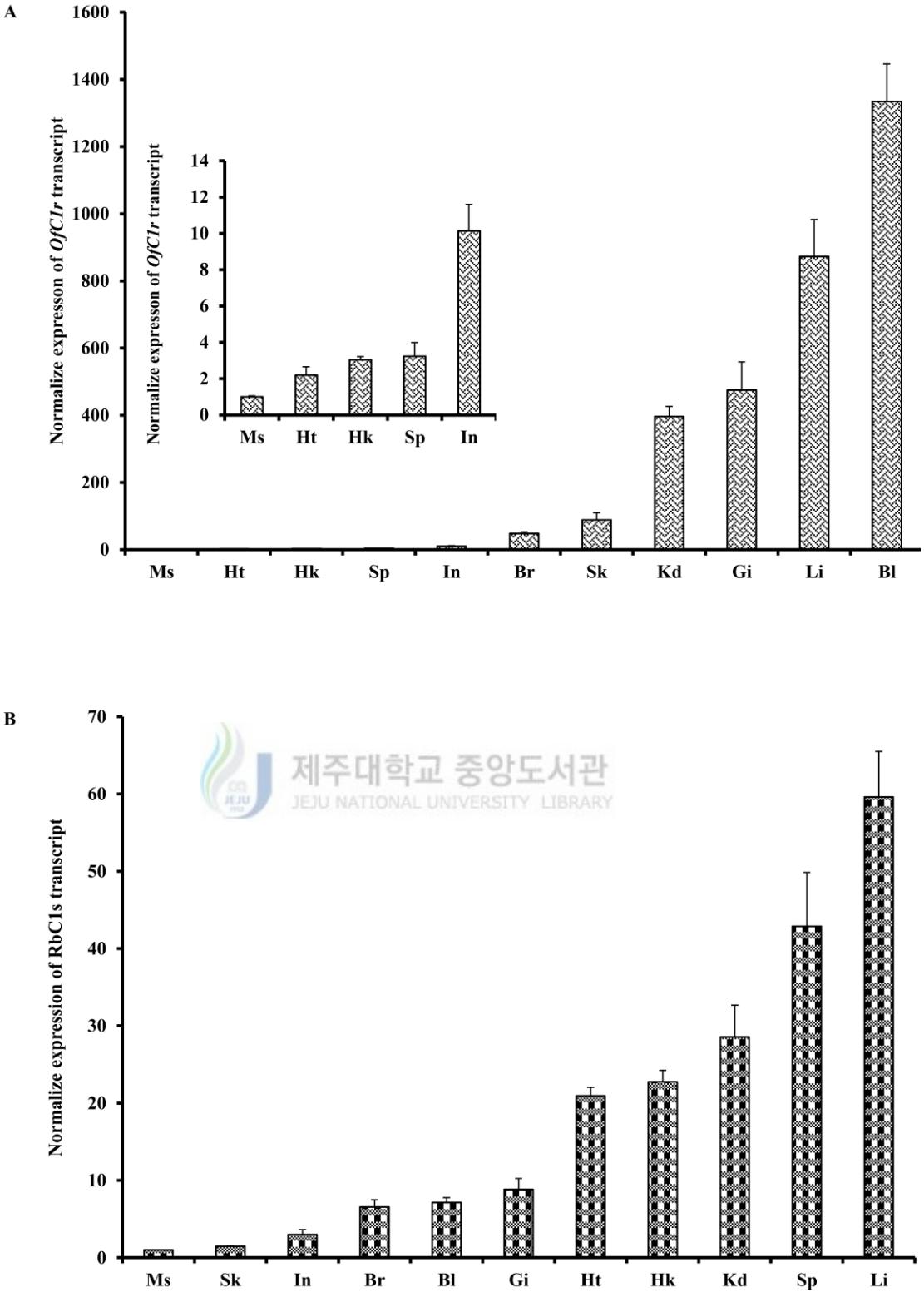
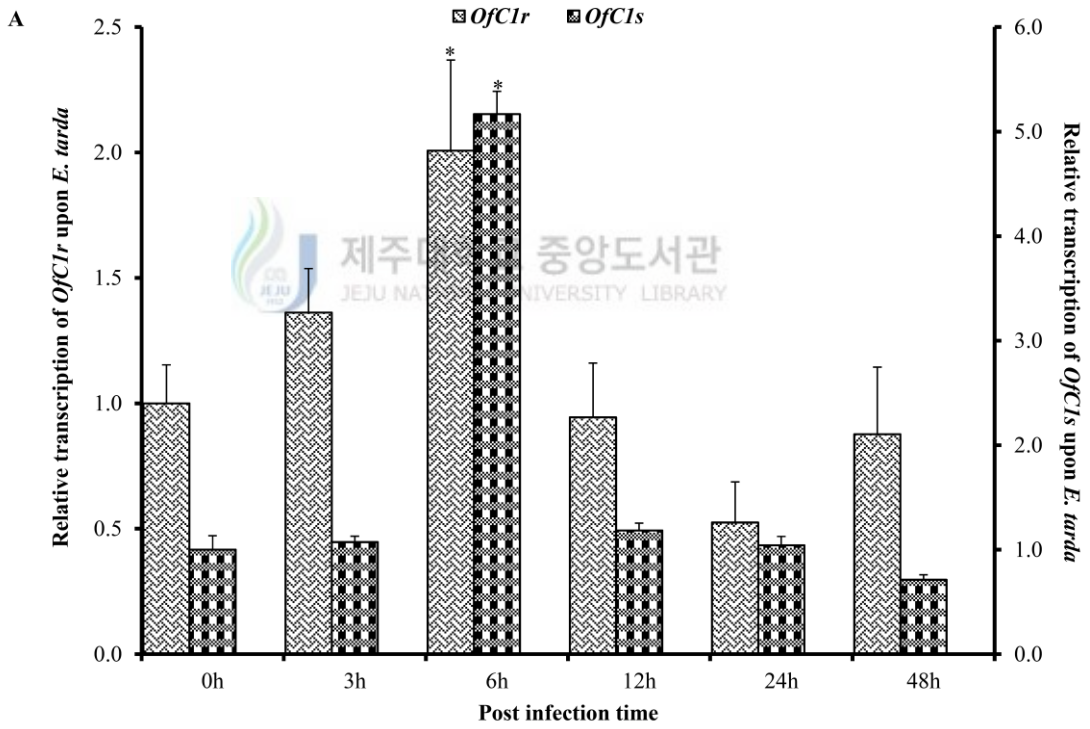


Figure 3.9. Tissue-specific mRNA expression analysis of OfC1r (A) and OfC1s (B) by qPCR. The mRNA expression in the peripheral blood cells (Bl), gills (Gi), liver (Li), spleen (Sp), head kidney (Hk), kidney (Kd), skin (Sk), muscle (Ms), heart (Ht), brain (Br), and intestine (In) was normalized using Ofβ-actin as the reference gene. The error bars represent the standard deviation of triplicate samples. Significant differences in expression were calculated with respect to expression in muscle for both *OfC1r* and *OfC1s* ($P < 0.05$).

3.2.6. Immune responsive transcription upon bacterial challenge

OfC1r and OfC1s mRNA expressions were analyzed in liver tissue isolated at different post infection time points from fish challenged with *E.tarda* and *S. iniae*. Upon *E.tarda* challenge, the highest mRNA expression of OfC1r and OfC1s were detected at 6 h p.i ($P<0.05$) in liver tissue (Fig. 3.10A). Similarly, OfC1s was shown greater mRNA expression at 6 h p.i upon the *S. iniae* infection, where OfC1r significantly up-regulated at early and later post infection of *S. iniae* (Fig. 3.10B).



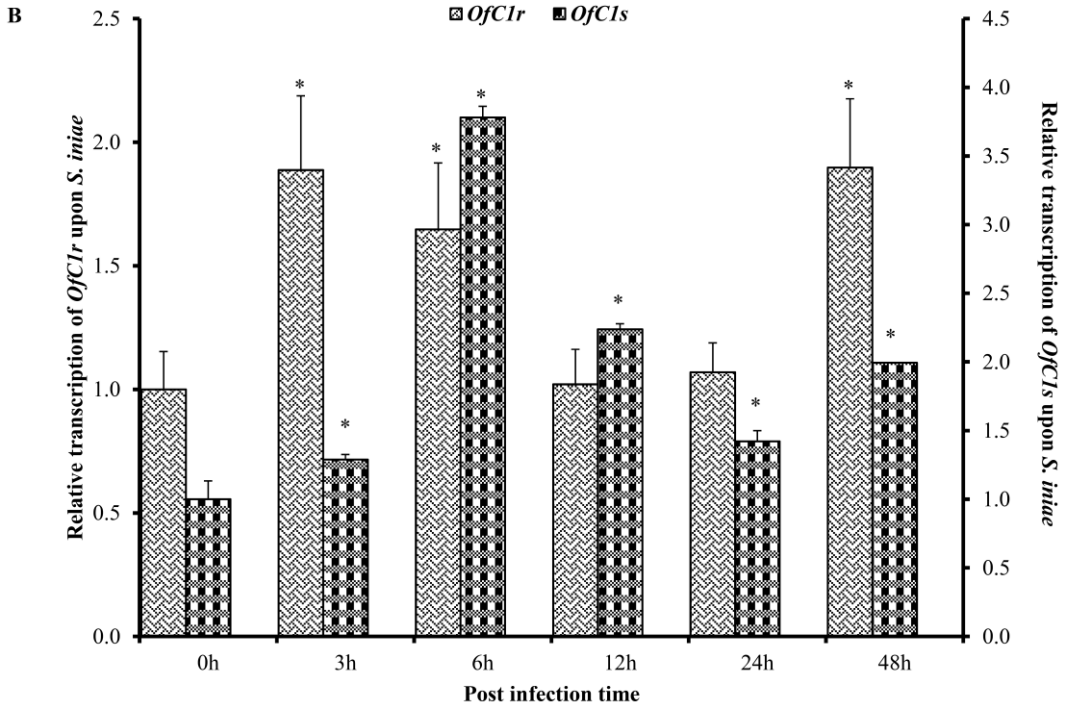


Figure 3.10. The temporal mRNA expression of OfC1r and OfC1s in liver tissue from rock bream challenged with *Edwardsiella tarda* (*E. tarda*) and *Streptococcus iniae* (*S. iniae*) and PBS (control). The mRNA levels were normalized to Ofβ-actin mRNA levels. Gene expression in each control group (0 h) was set at 1, and expression in each challenge group was expressed relative to the expression in the respective PBS control group (expression > 1: upregulated; expression < 1: downregulated). mRNA levels at different time points were compared using Student's *t*-test. Data from three individuals were analyzed. Vertical bars represent the mean ± SD (n = 3). Significant differences ($P < 0.05$), relative to the 0 h values, are indicated with an asterisk.

3.3. Discussion

First line immune responses are affirmed by numerous regulatory genes which cannot be compensated by the expression of another molecule. The C1r and C1s are two main subcomponents of the key recognition C1 complex in the classical complement pathway which provide first line defense against bacterial infections. In the present study, the molecular genomic level of OfC1r and OfC1s and, their mRNA expression profiles upon immune challenges were described. The C1r and C1s are highly specific serine proteases which are responsible for the catalytic function of C1 complex in complement system (Kardos et al., 2001). *In silico* characterization of

OfC1r and OfC1s were revealed the presence of serine protease domain at the C-terminal of their amino acid sequences. Generally, C1r and C1s are synthesized as zymogens and they need the proteolysis and cleave for them to become an active form. In addition, SPr domain is enriched with active sites and substrate binding sites which play a significant role in determining catalytic efficiency and substrate specificity. Mostly those sites were conserved among the other vertebrate counterparts of C1r and C1s, which confirmed by our multiple sequence analysis. Similar characteristic features of serine protease domain from C1r and C1s were previously found from human(Arlaud et al., 1998a; Carroll and Georgiou, 2013) and murine (Arlaud et al., 2002a; Garnier et al., 2003) that may suggesting similar in function. Furthermore, serine protease domain of C1r alone can cleave C1s with a rate comparable with that of the activation by the CCP1-CCP2-SPr catalytic fragment. Previously reported human C1r has been demonstrated that CCP region is critical for the activation and orientation of the C1r dimer (Kardos et al., 2008). The OfC1r homodimer is formed by CCP1-CCP2-SPr fragment in a head-to-tail fashion where the CCP1 module of one molecule interacts with the SPr domain of the other molecule and *vice versa*. This homodimeric structural arrangement of human C1r molecule was revealed the importance for the cleavage of C1s by forming the core of the C1s–C1r–C1r–C1s heterotetramer, which has to be exposed to the solution to carry out its function. Our tertiary modeling results are fully consistent with this picture of human (Arlaud et al., 2002b), may propose their involvement of proteolytic activity in rock bream.

Apart from above characteristic features, putative polypeptides of OfC1r and OfC1s were consisted with CUB and calcium binding EGF domain at the N-terminal region (CUB1-EGF-CUB2) where interaction is generated. The interaction of CUB-

EGF domain is significant for Ca^{2+}-dependent assembly of the CUB-EGF module pairs of C1r and C1s involved in the interactions. Previously, similar phenomenon had been demonstrated from human (Arlaud et al., 2002b; Thielens et al., 1999). This N-terminal region of both C1r and C1s proteins mediates the Ca^{2+} dependent C1r-C1s as well as C1q-$C1r_2$-$C1s_2$ association (Kardos et al., 2008; Villiers et al., 1985). This C1q-$C1r_2$-$C1s_2$ association may crucial in continue the complement classical cascade via recognizing the invading pathogens. Furthermore, (Arlaud et al., 2002a) report evidenced that C1r had ability to bind C1s in the presence of Ca^{2+}. More importantly EGF domain rich in calcium binding sites and EGF existence proteins require calcium for their biological functions as crucial for numerous protein-protein interactions, blood coagulation, fibrinolysis, neural development or cell adhesion (Arlaud et al., 2002a). In addition, each of the two C1r CUB modules bear N-linked oligosaccharide in the human protein (Arlaud et al., 2002a) whereas, OfC1r and OfC1s showed a single N-linked oligosaccharide in a single CUB domain. However, further studies are mandatory to clarify the OfC1r and OfC1s domain functions of rock bream fish. Thus, the presence of evolutionary conserved and functionally significant domains of OfC1r and OfC1s may represent similar function like mammalian orthologuess.

The results acquired by pairwise alignment, multiple alignment, evolutionary relatedness analysis and exon-intron structure comparison analysis exhibited the close proximity of OfC1r and OfC1s to the respective fish homologs. Specifically, multiple alignment results revealed the conserved active sites and substrate binding sites of OfC1r and OfC1s with human C1r and C1s, may suggest the similar catalytic function in serine protease domain (Arlaud et al., 1998a) throughout the evolution process.

In genomic structure analysis, the OfC1r and OfC1s genomes were comprised with ~6.5 kb and ~6.8 kb molecular weight, respectively. All the teleostean C1r

homologues have similar number of exons and on the other hand, higher vertebrate C1r consist with eleven exons. Interestingly, fish genomic sequences have been exhibited the collective lengths of last three exons of the CDS are almost similar in length of last CDS exon from the higher vertebrates. This may be due to the assembling of exons and removing of introns by RNA splicing during the evolutionary process. The genomic analysis of OfC1s was revealed similar number of exons in zebrafish, Mexican tetra and coelacanth. However, coelacanth C1s exon lengths are more close to the higher vertebrate C1s genome as its closer evolutionary proximity to the amphibian. On the other hand nine exons of mammalian C1s genomic sequences showed similar in lengths within the CDS. Genomic structures represent more reliable information regarding the evolutionary chronicle, since they undergo significant changes. Overall genomic findings may confirm the OfC1r and OfC1s genes have been evolved by a common ancestor of fish and mammals C1r and C1s, respectively. In addition, synteny analysis have been discovered that physical co-localization of both C1r and C1s genetic loci on the same chromosome within an individual. Amphibian and mammalians C1r gene duplication was observed throughout evolution process likewise OfC1r gene duplication evolving divergently has been detected. However, other fish species gene duplication was variable. Though, several other fish C1r/C1s sequences were deposited in GenBank, this is the first documentation describing a teleostean genomic structure arrangement and synteny analysis from rock bream in which OfC1r and OfC1s.

A ubiquitous expression of OfC1r and OfC1s transcripts were determined by spatial distribution analysis. The OfC1r mRNA was highly expressed in healthy blood and liver tissues where while OfC1s mRNA is mostly detected in liver and spleen. The C1q, C1r and C1s molecules are involved in the first line defense mechanism in

the complement system where C1q is synthesized in human myeloid cells, macrophages, monocytes, and dendritic cells (Rabs et al., 1986) and mouse C1r/C1s has a broad tissue distribution, with predominant expression in the liver as seen in other mammals (Garnier et al., 2003). In addition, teleostean spatial distribution pattern of the C1r mRNA has previously been determined in rainbow trout and depicted the highest expression in liver (Wang and Secombes, 2003) and also in common carp C1r/C1s mRNA mostly expressed in hepatopancrease (Nakao et al., 2001). Blood, liver and spleen are the immune tissues which are highly susceptible for many infectious microbes. Therefore, the spatial expression of OfC1r and OfC1s profiles may suggest the importance into immune responses as first line defense molecules. Similarly, the ubiquitous expression of OfC1r and OfC1s in all the examined tissues, including its non-immune tissues, indicated its primary importance, not only in immune regulation; but, also in other cellular functions which remains to be characterized.

Since, liver tissue is susceptible to infectious diseases the complement activation is prominent. Therefore, in order to examine the immune responses related to OfC1r and OfC1s, their mRNA levels in liver tissue were determined post-immune challenges. In order to activation of rock bream classical system, early phase up-regulation of rock bream C1q was reported upon *E. tarda* challenged liver tissue (Bathige et al., 2013). Obeying with classical cascade activation process, the OfC1r and OfC1s transcripts levels were up-regulated at early phase of *E. tarda* and *S. iniae* infection. As per our results, the *OfC1r* transcripts were expressed after *OfC1q* transcripts had been expressed (Bathige et al., 2013), may affirmed the activation of classical cascade system upon bacterial challenge. As the first documentation of C1r and C1s temporal expression from fish species, present study has been observed the

early phase immune responses upon bacterial infections. Furthermore, significant up-regulation of OfC1r and OfC1s transcripts in the immune related liver tissue may suggest potential involvement of rock bream innate immune responses.

In conclusion, the OfC1r and OfC1s genes were identified from rock bream and studied by analyzing the cDNA sequences, aa sequences, genomic sequences, tertiary structures and *in vivo* gene expression under normal and pathogenic infections. Collectively, all these findings suggest the involvement of OfC1r and OfC1s as a potential immune modulator in rock bream. Results of the current study further provide valuable insights into the temporal expression upon pathogenic infection of a vertebrate C1r and C1s from a teleostean species for the first time, and it may help to further explore the functional mechanisms of this vital gene in fish and other lineages to ultimately establish its biological roles.

4.0. Chapter II

Molecular characterization of complement factor D (CfD) from alternative pathway, as the convertor molecule of complement system

Molecular aspects, genomic arrangement and immune responsive mRNA expression profiles of complement factor D homolog from rock bream (*Oplegnathus fasciatus*)

Abstract

The complement factor D (CfD) is vital in alternative complement pathway activation via cleavage of the complement factor B. This catalytic reaction forms the alternative C3 convertase which is crucial for complement mediated pathogenesis. Rock bream (*Oplegnathus fasciatus*) CfD (OfCfD) have been completely characterized and their mRNA expression was investigated. Putative cDNA of OfCfD encodes a 239 aa with 26 kDa polypeptide. Domain analysis of OfCfD aa sequence discovered a single serine protease trypsin superfamily domain, three active sites and three substrate binding sites. Pairwise, homology sequence comparison displayed that OfCfD has highest identity (73.8 %) with the Nile tilapia CfD. Phylogenetic tree divided into two main clusters where one includes the fish CfD members and second cluster consisting with other vertebrate CfD members. The genomic structural arrangement of OfCfD (2451 bp) possessed five exons interrupted by four introns. The spatial transcriptional results indicated that OfCfD transcript was constitutively expressed in all examined rock bream tissues, as high in spleen and liver tissues. OfCfD transcripts were immunologically up-regulated by lipopolysaccharide (LPS) (12 h p.i.), *Streptococcus iniae* (12 h p.i.), rock bream iridovirus (RBIV) (6-12 h p.i.) and Poly: IC (6 h p.i.) in spleen tissue. Since, OfCfD is a trypsin protein, its recombinant protein has protease activity similar to trypsin (positive control), affirmed its catalytic function in alternative pathway. Together, our findings suggest that OfCfD has immune responsive functions in rock bream.

Keywords: Alternative pathway; genome; immune challenge; protease activity

4.1. Introduction

The serine proteases are a large family members of trypsin/chymotrypsin family proteins, functionally involved in diverse biological process, including complement activation, blood coagulation, fertilization, immunity and tissue development (Neurath, 1986). The serine protease complement factor D (CfD) is a member of trypsin family protein, functionally participate in alternative complement pathway activation (Volanakis and Narayana, 1996; Xu et al., 2001). The only known natural substrate of factor D is another complement serine protease, which is complement factor B (CfB). As a catalytic function of CfD polypeptide, it cleaves CfB into the fragments Ba and Bb.

The single Arg-Lys bond of factor B becomes susceptible to factor D-catalyzed hydrolysis only in the context of a Mg^{2+} dependent reversible complex between factor B and C3b, the major fragment of C3 activation (Muller-Eberhard and Gotze, 1972). In the alternative pathway of complement activation, the catalytic reaction leading to the formation of the alternative C3 convertase (C3bBb), which is responsible for cleavage of the complement protein C3. The cleaved C3 leads to the activation of terminal complement components to form the membrane attack complex on microbial or cellular surfaces (Harboe et al., 2004). Apart from that, CfD is involved in regulation of systemic energy balance and physiologic and pathologic processes, including immunity and inflammation (Rosen et al., 1989). The CfD is the only enzyme available in blood able to catalyze the C3 convertase formation reaction and is therefore absolutely required for alternative pathway activation (Lesavre and Muller-Eberhard, 1978). In contrast, the CfD is a rate-limiting enzyme in complement cascade because of its extremely low concentration in blood. Hence, it is considered as an important target for pharmaceutical control of complement activation.

Fish possess complement activation pathways similar to those in mammals, and the identified fish complement proteins have many homologies to their mammalian counterparts (Boshra et al., 2006; Holland and Lambris, 2002). However, much fewer studies have been conducted with CfD in teleost fish, particularly limited knowledge with its regulation after pathogenic infection. The molecular and functional properties of common carp CfD (Yano and Nakao, 1994), sequence properties and expression analysis of CfD from the channel catfish (Zhou et al., 2012) and olive flounder (Kong et al., 2009) were previously documented. Rock bream, *Oplegnathus fasciatus* is an economically significant fish species in Korea, which is under constant bacterial and viral threats. An enhanced knowledge on the defense mechanisms of rock bream against bacterial and viral infections will help us in developing methods to overcome infection and enhance disease free production of this valuable species. Hence, present study was conducted to (1) identify the putative cDNA and gDNA sequence and characterize its molecular and genomic structural features, (2) to evaluate the spatial mRNA expression profile in different tissues and temporal mRNA expression after immune challenges to demonstrate the possible involvement of CfD in the innate immunity of rock bream *Oplegnathus fasciatus*. The CfD molecule is designated as OfCfD.

4.2. Results

4.2.1. Sequence and domain structure depiction of OfCfD

The 1100 bp cDNA sequence of OfCfD was identified as an orthologue of a CfD member, via the BLASTX sequence comparison analysis. The OfCfD sequence information was deposited in GenBank under the accession No. KP241696. Putative nucleotide sequence of OfCfD harbor a 720 bp of coding sequence (CDS), which

could be coded a polypeptide of 239 amino acids with 26 kDa molecular mass and isoelectric point of 5.3. In the 3' untranslated region (UTR), poly(A) signal (^{980}AATAAA^{985}) is present (Fig. 4.1). Domain analysis of OfCfD protein sequence revealed that a prominent serine protease trypsin superfamily domain, three active sites (^{31}H, ^{78}D, ^{183}S) and three substrate binding sites (^{177}D, ^{198}S, ^{201}G).

```
                                                        AGTTAAACCGACATC    -15
ATGGAGTCAAAGAAA GATGTTCTCGTGGCT GCTGCTGTTTTTGTT GTTGCCCTCATTTCA     60
 M  E  S  K  K   D  V  L  V  A   A  A  V  F  V   V  A  L  I  S      20
TACAGTGAGGGCATA ATTGGTGGCAGAGAG GCAGCGCCACACTCT CGGCCCTACATGGCC    120
 Y  S  E  G  I   I  G  G  R  E   A  A  P  H  S   R  P  Y  M  A      40
TCCATCCAGGTGCCA GAAGGAGAGACTATG AAACATGAGTGTGGA GGGTTTGTGATTGCA    180
 S  I  Q  V  P   E  G  E  T  M   K  H  E  C  G   G  F  V  I  A      60
GATCAGTGGGTGATG ACTGCGGTACACTGT CTGCCGACAGGGCCA AATGGAAGGAAAGTG    240
 D  Q  W  V  M   T  A  V  H  C   L  P  T  G  P   N  G  R  K  V      80
GTGCTGGGTCTCCAT TCTCTGAGTGAACCT GAAGAAACAAAGCAT ACATTTGATATTTTG    300
 V  L  G  L  H   S  L  S  E  P   E  E  T  K  H   T  F  D  I  L     100
GAGCTTTACAATCAC CCAGACTTCAACCCG TCAAATTATGATAAT GACATTGCTTTAATT    360
 E  L  Y  N  H   P  D  F  N  P   S  N  Y  D  N   D  I  A  L  I     120
AAGTTGGATCGTCCA TTTAACACCTCTGAA GCTGTTAAAGCAGTG GAATACCTGCGAGCA    420
 K  L  D  R  P   F  N  T  S  E   A  V  K  A  V   E  Y  L  R  A     140
GGTGGCACTAACCCC GGCACAGGTGCAGAG GTCGAGACAGCCGGC TGGGGATCCCTCAAC    480
 G  G  T  N  P   G  T  G  A  E   V  E  T  A  G   W  G  S  L  N     160
GACCTGGGGTCCAGG CCAGACAAGCTCAAC GAGGTGGTCGTTGAA GTGTTCAGCTCGAAT    540
 D  L  G  S  R   P  D  K  L  N   E  V  V  V  E   V  F  S  S  N     180
CGGTGCCGACGCGGT GACTACTTTGGCTCA AAGTTCACCAGTAAC ATGATATGTGCACAT    600
 R  C  R  R  G   D  Y  F  G  S   K  F  T  S  N   M  I  C  A  H     200
AAAGTGTGCCCAGAC CCCTGCAATAAACCA CATAAGAAAGAAGAC AGTTGTGATGGTGAC    660
 K  V  C  P  D   P  C  N  K  P   H  K  K  E  D   S  C  D  G  D     220
TCTGGAGGTCCCCTG CTCTACAATGGCATT GCAGTGGGCATCACT TCCAATGGAGGAAAG    720
 S  G  G  P  L   L  Y  N  G  I   A  V  G  I  T   S  N  G  G  K     240
AAGTGTGGCCAAATA AAAAAGCCTGGAATT TACACTATCATCTCC CACTACACTGAGTGG    780
 K  C  G  Q  I   K  K  P  G  I   Y  T  I  I  S   H  Y  T  E  W     260
ATTGACAACACCATG GCCCTGCAGCCCGCT GCGACACAGGACCTG AGCAGTTAATAAAAT    840
 I  D  N  T  M   A  L  Q  P  A   A  T  Q  D  L   S  S  *           277
AAAAAGAGCAGGCAT CACAGTTACTAACTG CACATTACATTTTTC AGCAGAATTCACGTT    900
TCCACCACAGTCCTG CTTCTACTGAGTCAC TGATTATTCTACCTG GATATGAAGCTTATG    960
TGTTTGTGTGTTAGA CAGAAGTGTTCCTTT CTCATTTGCCACTAA TGCATGCCTCTAGAG   1020
ACAGATCTCCGGGAT GGCGAACATATTAGC CAGAATAATATATTT CACATAATTTTCACT   1080
TTATAAAAGAATCAA TAAAATACACTTGAT C                                 1111
```

Figure. 4.1. Nucleotide and deduced amino acid sequences of OfCfD. The nucleotide sequence (upper) and the deduced amino acid sequence (lower) are numbered. The start codon (ATG) and stop codon (TAA) are bolded and boxed. The serine protease trypsin family domain is grey shaded. Active sites in the trypsin domain are bold and underlined. The substrate binding sites are circled.

4.2.2. Homology sequence comparison and phylogenetic analysis

The pairwise homology assessment of OfCfD with other fifteen CfD members revealed that OfCfD shared highest identity (73.8 %) and similarity (78.9 %) with the Nile tilapia CfD homologue (Table 4.1). Second highest identity was recognized with zebra mbuna (*Maylandia zebra*) fish with 70.9 %. The fish CfD homologues showed the more than 50 % identity with OfCfD amino acid sequence. Other vertebrate counterparts indicated less than 40 % identity with CfD homologue.

Table 4.1. Percent identity and similarity for the amino acid sequences of OfCfD from different organisms

	CfD	Identity %													
		O. fasciatus	*O. niloticus*	*M. zebra*	*P. olivaceus*	*T. rubripes*	*X. maculatus*	*I. punctatus*	*D. rerio*	*O. orca*	*H. sapiens*	*B. taurus*	*S. scrofa*	*G. gallus*	*R. norvegicus*
Similarity %	*Oplegnathus fasciatus*		73.8	70.9	69.7	68.1	66.8	52.0	50.4	34.1	32.7	31.9	31.8	31.7	28.9
	Oreochromis niloticus	78.9		96.0	79.8	76.8	75.8	55.4	53.3	38.6	37.5	36.4	27.3	35.8	33.9
	Maylandia zebra	78.9	98.9		77.3	75.4	75.8	55.1	52.2	39.3	37.9	37.1	27.3	36.4	33.9
	Paralichthys olivaceus	78.3	88.1	87.0		72.9	70.8	54.9	54.5	40.2	38.4	36.3	26.7	37.3	34.6
	Takifugu rubripes	77.5	86.2	85.5	84.1		73.6	55.1	52.5	41.1	38.4	37.9	28.3	36.6	33.5
	Xiphophorus maculatus	76.4	87.6	87.6	84.8	85.9		57.1	55.4	38.1	35.7	36.3	26.5	37.0	33.7
	Ictalurus punctatus	69.2	70.5	70.2	70.0	70.7	71.6		68.3	39.8	39.3	40.2	29.1	41.9	37.7
	Danio rerio	66.7	69.5	68.0	70.0	67.8	68.7	78.8		41.0	41.4	41.4	27.2	36.8	38.6
	Orcinus orca	52.9	57.8	58.2	57.8	60.1	58.2	60.2	61.8		80.3	88.8	64.5	62.8	66.3
	Homo sapiens	53.1	55.6	56.0	54.9	55.4	55.6	59.6	62.7	85.3		76.4	64.1	58.5	64.3
	Bos taurus	51.4	56.0	56.4	56.0	58.3	57.1	59.5	60.6	93.4	84.6		65.3	62.8	64.8
	Sus scrofa	45.6	39.3	40.0	39.4	41.3	40.4	44.4	45.8	68.3	70.4	69.1		45.7	47.9
	Gallus gallus	47.2	54.9	56.4	55.6	55.8	57.5	60.0	57.0	73.6	67.9	74.7	54.0		53.8
	Rattus norvegicus	46.8	54.9	54.5	52.0	52.9	52.0	56.3	57.0	79.8	76.8	79.5	58.9	70.2	

Multiple sequence alignment of CfD from different organisms exhibited conservation in the trypsin superfamily domain, especially within the fish members (Fig. 4.2). Other than that, the three active sites were highly conserved among the selected vertebrate CfD homologues and three substrate binding sites were conserved within the fish CfD counterparts.

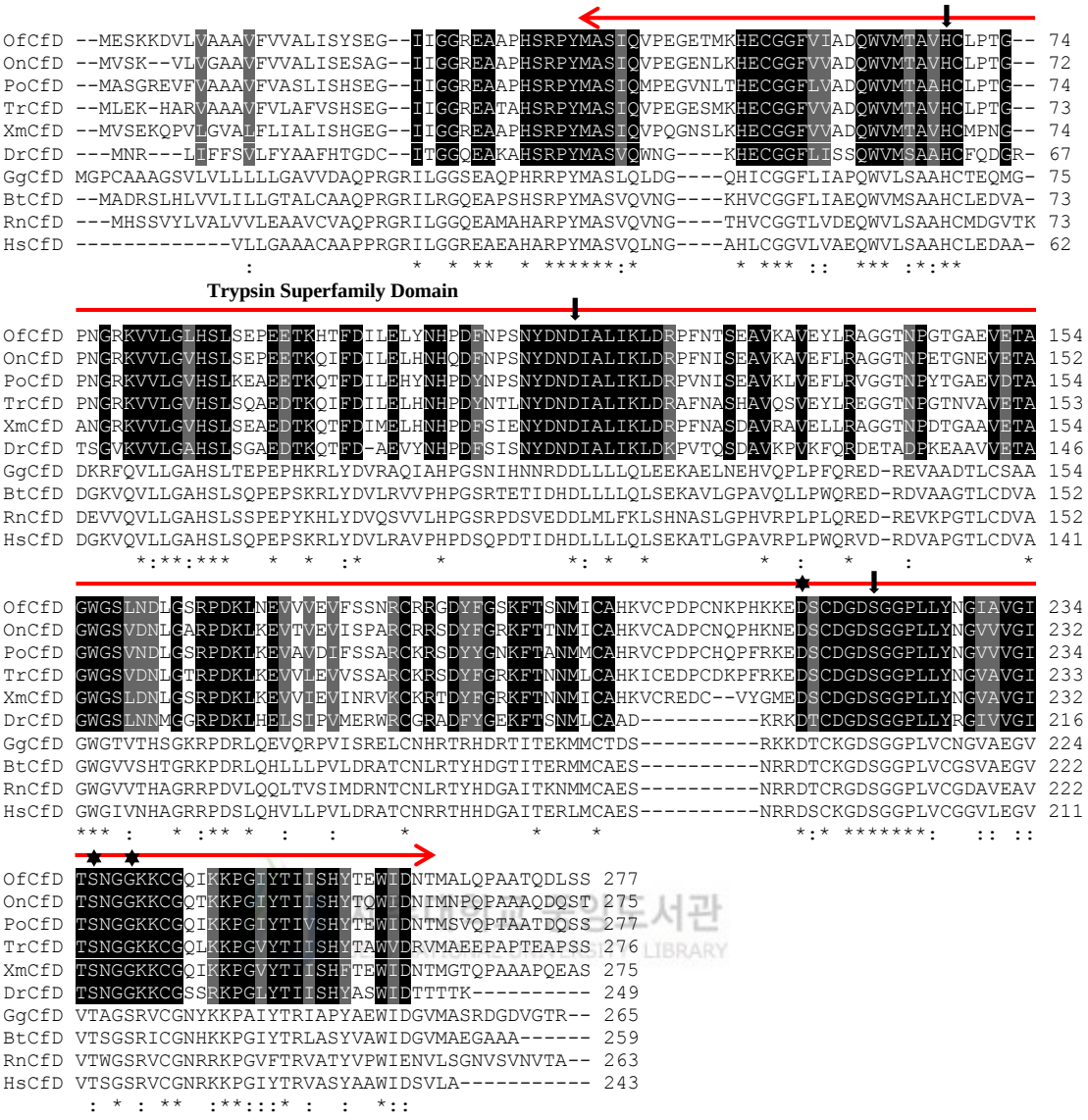


Figure 4.2. Multiple sequence alignment of OfCfD with other known CfD aa sequences using ClustalW (v2.1). Identical residues among selected organisms are denoted by (*) and similar residues are denoted by (:). Missing amino acids are indicated by dashes. Conserved residues within the fish counterparts were black shaded and similar residues were grey shaded. The serine protease trypsin superfamily domain is marked in red color double head arrow. Active sites were marked by black color downward arrow and substrate binding sites were marked by black shaded star mark.

The phylogenetic analysis was conducted to understand the evolutionary conservation of CfD gene in organisms by reconstructing a NJ-tree based on amino acid sequence alignment of OfCfD and those from other biological origins (Fig. 4.3). The results were revealed two main clusters, one including the fish CfD members and

second cluster consisting with other vertebrate CfD members. OfCfD clustered among the fish CfD group, particularly closer to the flounder homolog.

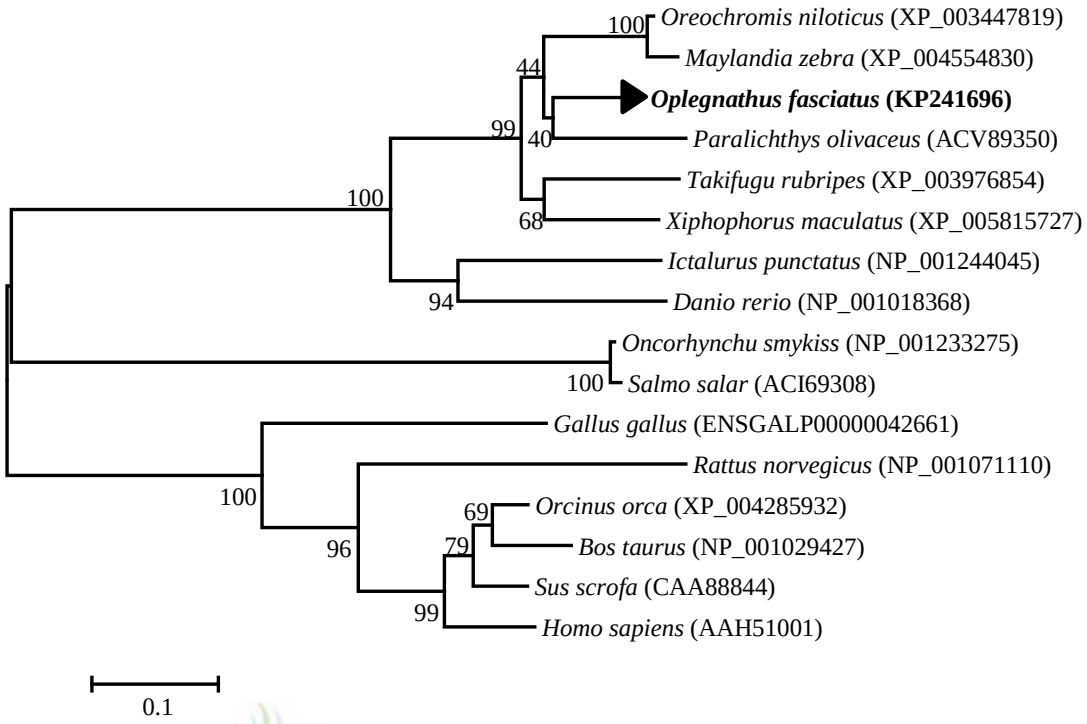


Figure. 4.3. Phylogenetic analysis of OfCfD with their known CfD orthologues. The tree is based on the alignment of the full-length amino acid sequences. The tree was constructed using ClustalW (v2.1) and MEGA (v5.0). The numbers at the branches denote the bootstrap replications for 5000 replicates. The GenBank accession numbers are shown within brackets next to each species.

4.2.3. Illustration of OfCfD tertiary structure

The features that make OfCfD have been illustrated based on the X-ray crystal structure of a human serine protease trypsin super family domain. We adopted this partial crystal structure as template (PDB ID, 4d9q) to construct the model of OfCfD (Fig. 4.4). The trypsin super family domain of OfCfD had fifteen β-sheets and four α-helices. Apart from that, three active sites and three substrate binding sites were scattered close to each other.

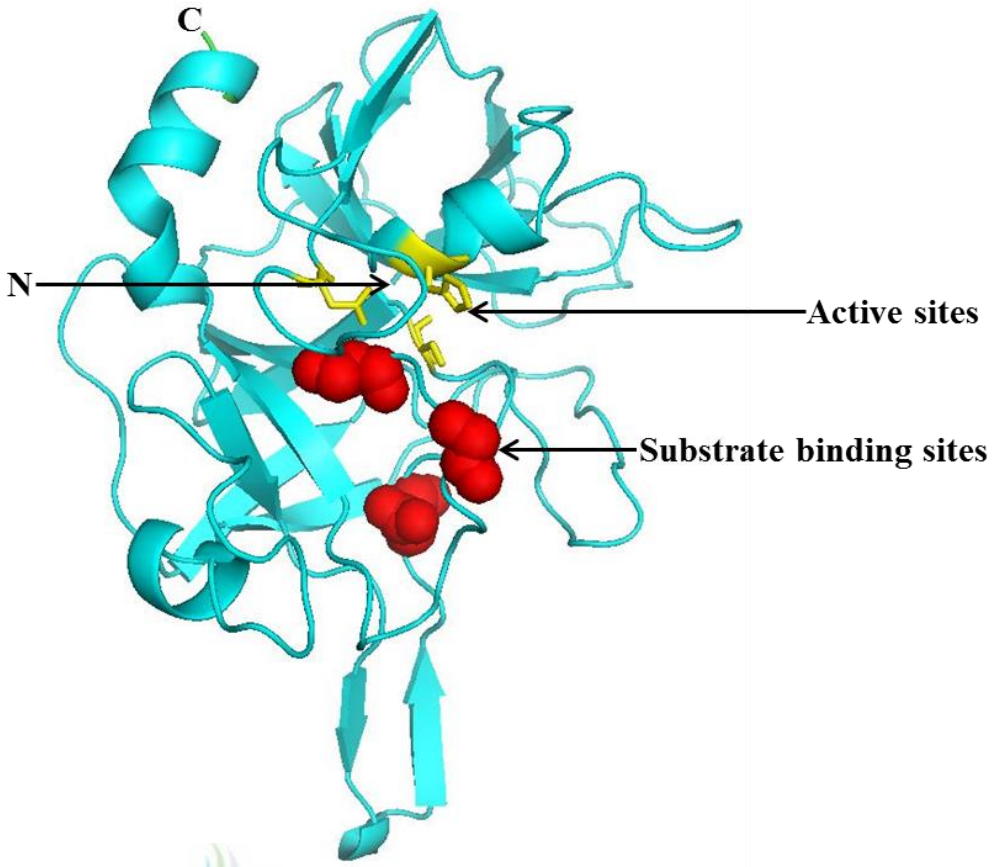


Figure. 4.4. Comparative tertiary structure of OfCfD putative protein molecule. The putative serine protease trypsin superfamily domain is blue colored. The active sites are marked in yellow and substrate binding sites are marked in red color spheres.

4.2.4. Comparative genomic structural analysis of OfCfD homologues

The genomic sequence of the OfCfD was 2451 bp and possessed the five exons interrupted by four introns (Fig. 4.5). The exon intron junctions were consistent with the GT/AG splicing rule. Two fish homologs, Nile tilapia and fugufish CfD, shared three similar exon lengths with the OfCfD within the CDS. However, OfCfD 5'UTR was divided into two exons where other homologue consisted with single exon in 5'UTR. Moreover, two exons from mammalian counterparts were shown length similarity among them. However, the genomic structures of CfD members from different taxa were shown significant variance in the context of genome size though, presence of similar number of exons within the selected CfD homologues.

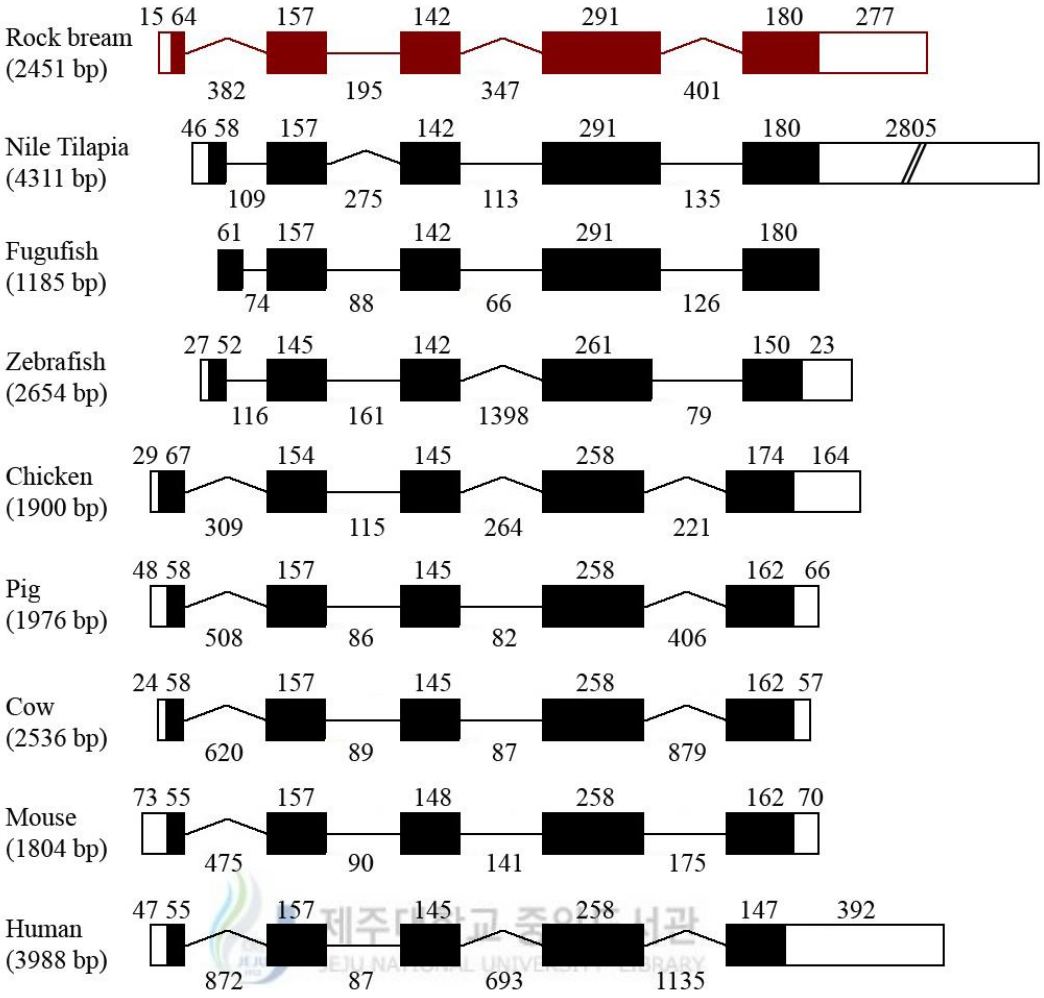


Figure 4.5. Schematic genomic structures of CfD (rock bream: KP241696, Nile tilapia: ENSONIT00000015561, fugufish: ENSTRUT00000012080, zebrafish: ENSDART00000057825, chicken: ENSGALT00000044292, pig: ENSSSCT00000014662, cow: ENSBTAT00000063284, mouse: ENSMUST00000061653, human: ENST00000327726) from various vertebrates. Coding regions are shown as dark shaded boxes. The 5′- and 3′- untranslated regions are denoted with empty boxes and introns are indicated by lines. The sizes of the exons are shown above the exon boxes, and the intron sizes are shown below the intron lines.

4.2.5. Spatial transcriptional profile of OfCfD in healthy rock breams

Spatial expression analysis of OfCfD in normal tissues revealed ubiquitous presence, with predominant expression in spleen (Fold-611.27), followed by liver and kidney. Head kidney, heart, intestine and gill were shown OfCfD transcription at moderate level where, muscle exhibited marginal expression of OfCfD transcript (Fig. 4.6).

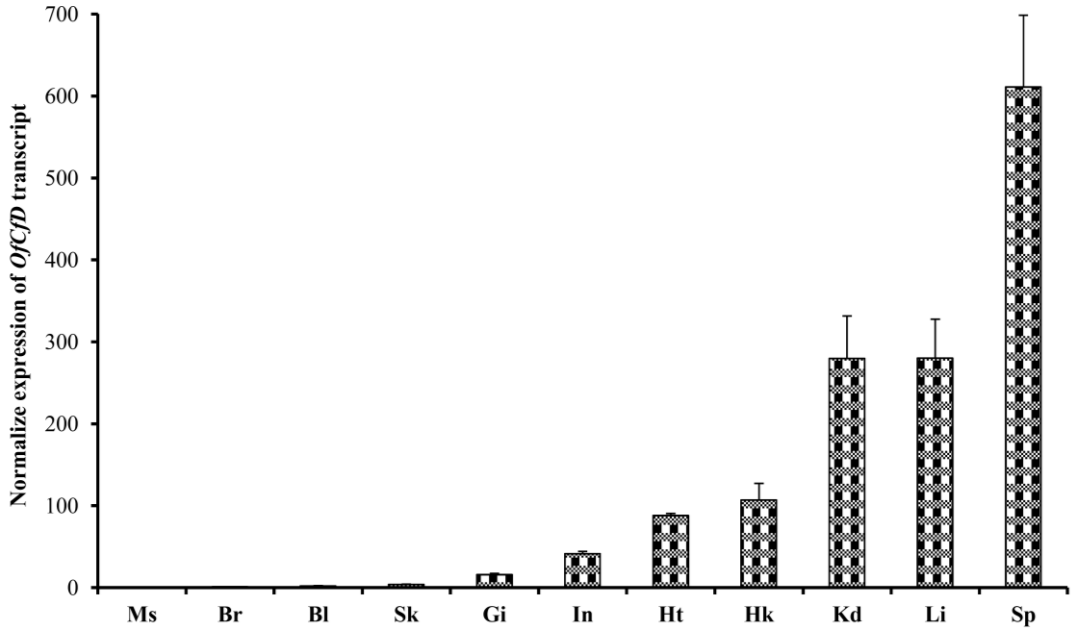


Figure 4.6. Tissue-specific mRNA expression analysis of OfCfD by qPCR. The mRNA expression in the peripheral blood cells (Bl), gills (Gi), liver (Li), spleen (Sp), head kidney (Hk), kidney (Kd), skin (Sk), muscle (Ms), heart (Ht), brain (Br), and intestine (In) was normalized using Ofβ-actin as the reference gene. The error bars represent the standard deviation of triplicate samples. Significant differences in expression were calculated with respect to expression in muscle ($P < 0.05$).

4.2.6. Immune responsive transcription upon bacterial and viral challenge

Temporal mRNA expression of OfCfD was analyzed in spleen tissue isolated from fish challenged with LPS, *S. iniae*, Poly: IC and RBIV (Fig. 4.7). As a response of pathogenic bacterial infection, the *S. iniae* and LPS challenges were displayed similar up-regulation pattern of OfCfD transcripts with maximum level of expression at 12 h p.i ($p<0.05$). As an immune responsive transcriptional regulation upon viral infections, the RBIV challenged fish showed significant up-regulation of OfCfD transcripts at 6 h and 12 h p.i, whereas Poly: IC injected fish outstandingly expressed the OfCfD transcripts at 6 h p.i.

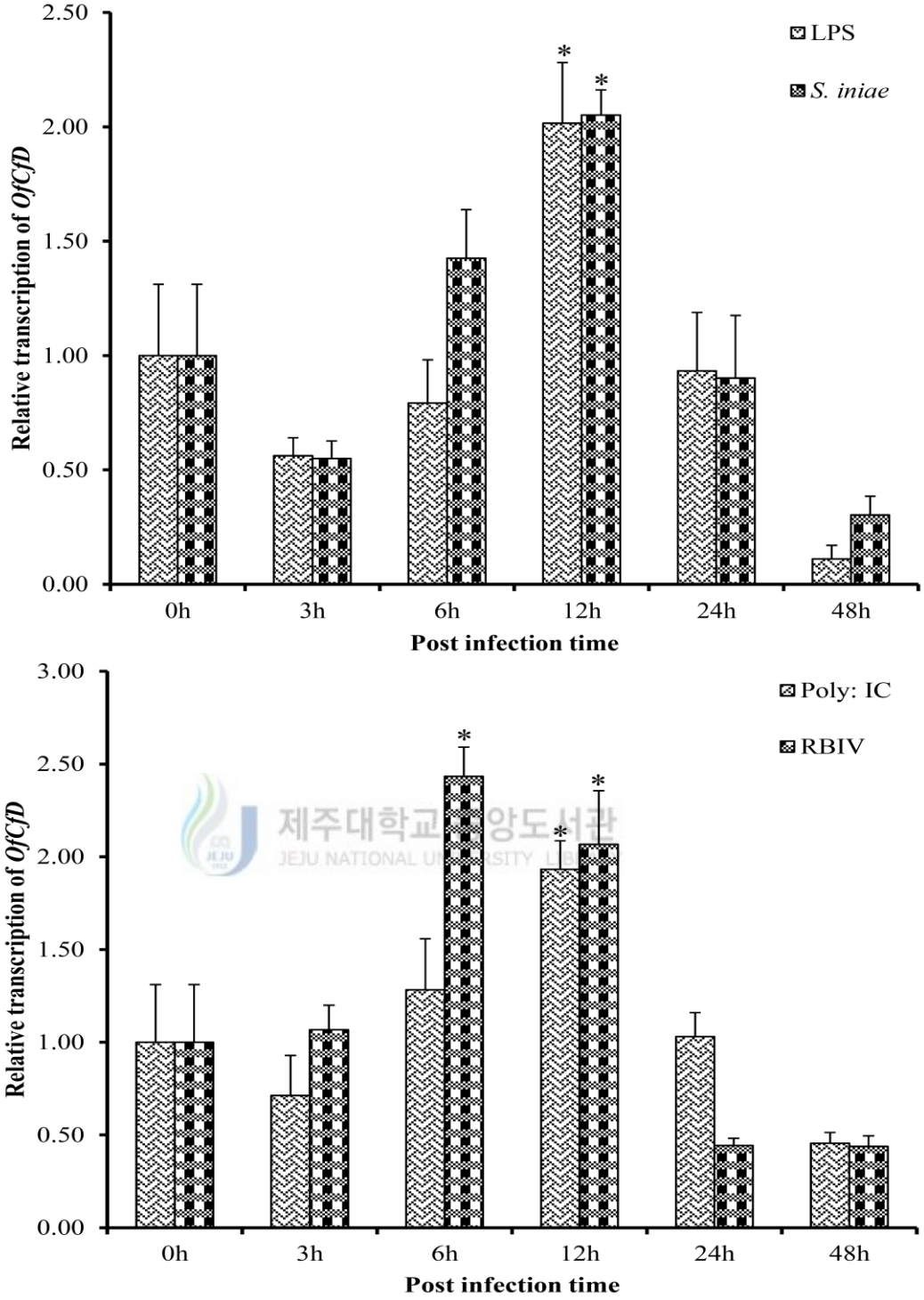


Figure 4.7. The temporal mRNA expression of OfCfD in spleen tissue from rock bream challenged with LPS, *S. iniae*, RBIV, Poly: IC and PBS (control). The mRNA levels were normalized to Ofβ-actin mRNA levels. Gene expression in each control group (0 h) was set at 1, and expression in each challenge group was expressed relative to the expression in the respective PBS control group (expression > 1: upregulated; expression < 1: downregulated). mRNA levels at different time points were compared using Student's *t*-test. Data from three individuals were analyzed. Vertical bars represent the mean ± SD (n = 3). Significant differences ($P < 0.05$), relative to the 0 h values, are indicated with an asterisk.

4.2.7. Protease activity

SDS-PAGE analysis revealed the expression of recombinant OfCfD fusion protein (OfCfD: 26 kDa + MBP: 42.5 kDa = rOfCfD+MBP: 68.5 kDa) (Fig. 4.8).

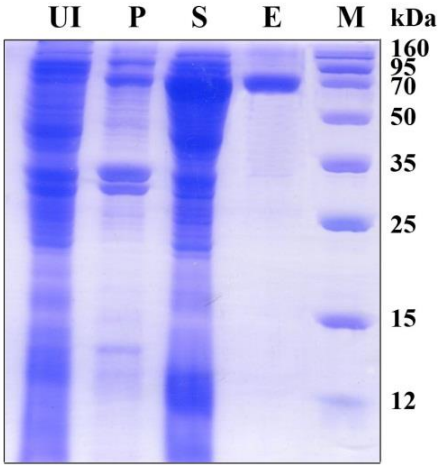


Figure 4.8. SDS-PAGE analysis of recombinant OfCfD protein. UI: un induced, P: pellet, S: Supernatant after pelleted, E: purified protein elution and M: protein marker.

The significant protease activity was shown by recombinant OfCfD protein to the negative control which is recombinant MBP. In addition, rOfCfD showed almost similar protease activity to the trypsin which is positive control (Fig. 4.9).

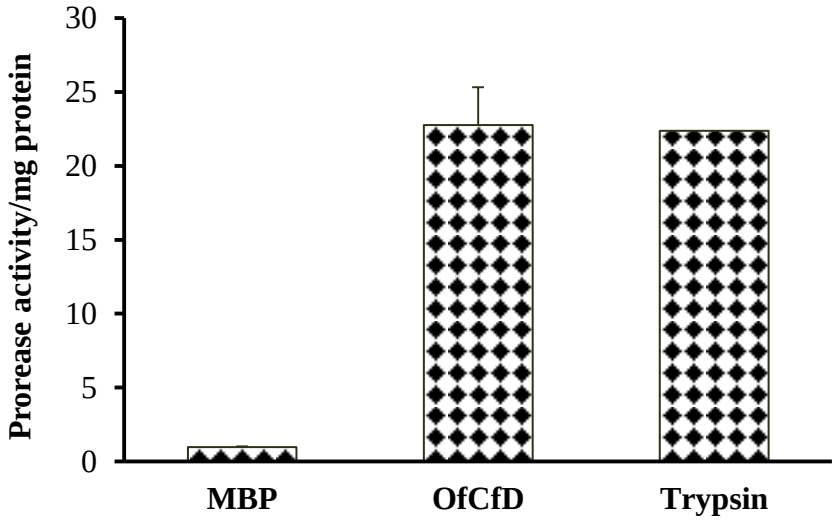


Figure 4.9. Protease activity of OfCfD recombinant protein. MBP: recombinant Maltose Binding protein (negative control), OfCfD: recombinant OfCfD protein, Trypsin: Positive control.

4.3. Discussion

All organisms are exposed to a challenging environment comprised of various pathogenic infections. Inflammation and acute phase response are components of the innate immune response against infection to protect the organisms (Baumann and Gauldie, 1994). A gene involved in alternative complement pathway which is complement factor D, catalyzing the enzymatic reaction leading to the formation of the C3-convertas to regulate the immune and inflammatory processes via complement activation (Volanakis and Narayana, 1996). The putative cDNA and gDNA sequences of CfD were identified from rock bream (OfCfD) and characterize in terms of its sequences structure and genomic organization. Further OfCfD spatial mRNA expression profile in normal tissues and expression profile after bacterial and viral infections were analyzed to discover its innate immune responses.

Molecular characterization of OfCfD revealed a CDS encoding for a protein of 239 aa, similar to that of *Danio rerio* CfD (249 aa) and *Homo sapiens* CfD (243 aa) but different from the other selected teleostean (275-277 aa) and mammalian (259-263 aa) homologues. However, pairwise sequence comparison of OfCfD was exhibited more than 50 % amino acid sequence identity with fish CfD sequences where higher vertebrate CfD showed less than 35 % identity. On the other hand, OfCfD consisted with a single trypsin superfamily domain ranged in the CDS and our multiple sequence analysis results revealed its high conservation residues within the teleostean CfD counterpart. Interestingly, complement factor D is structurally differs from all other complement enzymes, in that it has a single serine protease domain with no additional modules or domains (Arlaud et al., 1998b). Therefore, it diverges from all other serine proteases in the unique mechanism regulating its proteolytic activity. In order to accomplish the OfCfD proteolytic activity, it bears three

catalytically active sites and three substrate binding sites where located within the trypsin superfamily domain. As only known substrate of the CfD, the serine protease CfB is hydrolyzed at the Arg-Lys bond (Volanakis and Narayana, 1996) via bind to substrate binding sites of CfD. Our multiple sequence analysis was revealed the highly conserved active sites and substrate binding sites among the vertebrate CfD homologues, affirming their significance of OfCfD on its proteolytic activity. Furthermore, conservation of the domain features and arrangements in the consensus aa sequence of OfCfD indicated that OfCfD encoded a functional homolog of CfD, and further suggesting similar functions described for mammalian CfD (Jing et al., 1998). Phylogenetic analysis also indicated the distinctive nature of CfD sequences, as they formed a separate two main clusters where one consisted with fish counterpart and second cluster with aves and mammals. In addition, the *Oncorhynchu smykis* CfD and *Salmo salar* CfD have been arisen in a different rate of evolution compare to other fish CfD members, as revealed by their clustering arrangement. The reconstruction of CfD phylogeny and its clustering pattern clearly revealed that OfCfD has common ancestral origin through a different rate of their evolution process and confirmed the clear divergence from teleost CfDs.

In order to establishing the aspects of evolution and roles of CfD, one of the important steps is to determine its genomic arrangement. The genomic organization of CfD has been reported for only in channel catfish (Zhou et al., 2012). We were able to determine and compare the genomic organization of OfCfD with that of other counterparts from different lineages, and to deduce any structural changes occurred in CfD gene structure during the evolution. Similar to the other teleost, aves and mammalian CfDs, the OfCfD depicted five exons interrupted by four introns. The OfCfD genomic arrangement is appeared to be distinct from the teleost counterpart in

terms of the number of exons in the CDS where OfCfD consist four exons and others consisted of five exons in their CDSs. However, these genomic evidences indicate that, it has maintained an overall integrity throughout the evolution. Our data clearly demonstrated the evolutionary conservation of CfD in terms of its sequences, genomic arrangement and tertiary structure.

Ubiquitous spatial expression of OfCfD was observed in juvenile rock bream tissues with a most expression in spleen and liver. In addition, teleostean spatial distribution pattern of the CfD mRNA has previously been determined in channel catfish and depicted the highest in gill and spleen (Zhou et al., 2012) and also flounder possessed abundant CfD expression in gill, heart, liver and spleen (Kong et al., 2009). However, CfD expression pattern is differed from that observed in human (White et al., 1992). The spleen is a lymphatic organ that plays a fundamental role in protecting the body from invading pathogens (Tiron and Vasilescu, 2008). The liver plays a role in innate immunity upon infection by pathogens through the synthesis and secretion of acute phase proteins, such as pro-inflammatory cytokines. The gills are one of the first lines of defense against infection, representing the main interface between the fish and an environment full of pathogens (Lichtenfels et al., 1996). Therefore, the constitutive expression of OfCfD in spleen, liver, kidney and head kidney may imply a crucial role for OfCfD in the rock bream innate immune defense.

To investigate the immune responsive expression level of OfCfD mRNA was affected through viral or bacterial infection, we have been employed various host-pathogen models. LPS and *S. iniae* used as bacterial immune stimulant where LPS is a major component found in the outer membrane of Gram-negative bacteria, and elicit strong immune responses in animals. Therefore, it is routinely used to study the expression pattern of immune-related genes. In this study, LPS and *S. iniae*

challenged rock bream spleen revealed significant OfCfD expression at 12 h p.i. Previously, channel catfish has been reported the CfD gene expression (at 24 h p.i.) in spleen tissues, after challenge with *Edwardsiella ictaluri* (Zhou et al., 2012). Also, *S. iniae* challenged flounder spleen and liver has shown up-regulated CfD mRNA expression at 1 h p.i. and 12 h p.i., respectively (Kong et al., 2009). In addition, RBIV and Poly: IC was used as a viral immune stimulant where rock bream iridovirus is widely distributed disease and is responsible for massive losses in the rock bream aquaculture (Li et al., 2011). Upon infection of RBIV and Poly: IC, the expression of OfCfD mRNA in spleen increased at 6 h p.i. (up to 12 h p.i.) and 12 h p.i., respectively. In a similarly context, flounder (Kong et al., 2009) has been reported the CfD expression is increased at 12 h post infection of viral hemorrhagic septicemia virus (VHSV) in both spleen and liver tissues. Collectively, a transcriptional result of OfCfD is significantly up-regulated at early phase after the bacterial and viral challenge may suggest the innate immune responses of OfCfD in rock bream. Our data, and the previous results, provide evidence for the potential innate immune role for OfCfD in rock bream against bacterial and viral infections. Furthermore, results may suggest that the expression of complement factor D homologs is modulated by complex mechanism(s) in the host defense system.

As a trypsin family serine protease, rOfCfD could cleave the casein substrate while depicting similar protease activity to the universal trypsin protein. Therefore, present results were affirmed that rOfCfD is a trypsin family protease member. This phenomenon further correlated with the domain architecture analysis results, in which OfCfD consisted with trypsin superfamily domain. On the other hand, fusion protein tag which is rMBP does not possess protease activity. These results may suggest the OfCfD is a functionally active serine protease involved in alternative cascade system.

In conclusion, we have characterized the CfD ortholog of rock bream (OfCfD) with the aid of molecular features, genomic arrangements, immune responsive transcriptional modulation and its protease activity. OfCfD shared several structural characteristics with other CfD orthologous. Ubiquitous OfCfD expression with higher levels in spleen and liver was detected by qPCR. The transcriptional modulation under pathological conditions indicated possible alternative complement activation upon antibacterial and antiviral defense.

5.0. Chapter III

Molecular characterization of complement carboxypeptidase N small subunit (CPN1) as the anaphylatoxin inactivator of complement system

Carboxypeptidase N small subunit (CPN1) from rock bream (*Oplegnathus fasciatus*): Molecular characterization, genomic arrangement and immune responses upon pathogenic stresses

Abstract

Carboxypeptidases (CPs) are proteases that hydrolyze C-terminal peptide bonds. They are involved in regulating the complement system of the immune system. Here, we report the molecular characterization and immune response of carboxypeptidase N1 from rock bream (OfCPN1). The complete cDNA (1591 bp) possessed 1347 bp of a coding sequence (CDS) (including stop codon) encodes a putative polypeptide with 448 aa. Predicted molecular weight and isoelectric point was 51 kDa and 6.5, respectively. The amino acid sequence of OfCPN1 was found to harbor a signal peptide (19 aa), two characteristic Zn-binding signature domains, active Zn binding sites and single peptidase-M14 Zn carboxypeptidase site. Pairwise analysis revealed that OfCPN1 had the highest identity (96.9%) and similarity (97.8%) with the *Dicentrarchus labrax*. The genomic sequence of OfCPN1 has 9 exons interrupted by 8 introns. qPCR results indicated that OfCPN1 was constitutively expressed mainly in the liver, kidney, heart and head kidney. OfCPN1 was transcriptionally regulated in the liver upon challenge with pathogenic bacteria (*Edwardsiella tarda*, *Streptococcus iniae*), RBIV and the immune modulators, Poly: IC and LPS. Taken together, our findings suggest that OfCPN1 have immune-related functions in rock bream.

Keywords: anaphylatoxin; carboxypeptidases; genome; immune challenge

5.1. Introduction

Carboxypeptidases (CPs) are a group of protease enzymes that hydrolyze C-terminal peptide bonds. They are generally classified into several families based on their active site mechanism and sequence similarities. According to the active site, serine CPs (CPY/C, CPD) (Remington and Breddam, 1994), cysteine CPs (lysosomal CPB) (Cammack, 2006) and metallo CPs. Metallocarboxypeptidases are further divide into two sub groups based on sequence similarity. Two groups were carboxypeptidase A/B (CPA, mc-CPA, CPB, CPU/R) and carboxypeptidase N/E (CPN, CPH/E, CPM, CPZ). Both metallocarboxypeptidase subgroups have Zn as a metal ion; therefore, CPA/B and CPN/E are also referred to as Zn metallocarboxypeptidases (Keil et al., 2007; Pejler et al., 2009). Furthermore, based on their substrate specificity, Zn metallocarboxypeptidases are sorted into two subgroups: CPA/B, whose members cleave hydrophobic C-terminal amino acids, and CPN/E, whose members cleave C-terminal arginine and lysine (Matthews et al., 2004; Pejler et al., 2009).

CPN is also known as kininase I, arginine carboxypeptidase, or anaphylatoxin inactivator (Davis et al., 2005). Structurally, CPN (EC: 3.4.17.3) comprises two small subunits (CPN1) ranging in size from 48 to 55 kDa and two large non-catalytic subunits (CPN2) (83 kDa) that form a 280-kDa tetramer (Keil et al., 2007; Levin et al., 1982; Plummer and Hurwitz, 1978; Skidgel and Erdos, 2007). CPN1 regulates the biological activities of C3a C4a and C5a in the complement system to maintain tissue homeostasis and protect tissues from anaphylatoxin-induced effects, that is, degranulation, smooth muscle contraction, chemotaxis, and increase in blood capillary permeability (Levin et al., 1982; Matthews et al., 2004; Skidgel, 1995). Briefly, CPN1 cleaves the C-terminal arginine of C3a to produce the derivative C3a-desArg

(Bokisch and Muller-Eberhard, 1970; Matthews et al., 2004), which does not bind the C3a receptor or mediate the pro-inflammatory activities attributed to C3a, C4a, and C5a (Hsu et al., 1997; Mathews et al., 1986; Matthews et al., 2004; Matthews and Wetsel, 2001; Wetsel et al., 2000). Moreover, CPN1 can regulate peptide activity and receptor binding by removing a single amino acid at the C terminal (Bokisch and Muller-Eberhard, 1970; Erdos and Sloane, 1962; Matthews et al., 2004; Matthews and Wetsel, 2001; Wetsel et al., 2000) and can completely abolish the anaphylatoxin activity of C3a, C4a, and C5a (Bokisch and Muller-Eberhard, 1970).

Although the CPN gene from several vertebrate organisms have been identified and characterized. However, this is the first characterization and documentation of CPN1 gene from teleosts. Therefore, it will be useful to characterize the CPN1 gene and their immune-related responses in other teleosts. Furthermore, studying of immune-related genes could facilitate the development of disease control strategies for rock bream. In this study, we isolated the OfCPN1 sequences from a rock bream cDNA and BAC library (Umasuthan et al., 2013). The cDNA and genomic DNA sequences, phylogenetic relationships, and tissue-specific mRNA expression were characterized, and the immune responses of OfCPN1 to bacterial, viral, and immune modulator challenges were assessed.

5.2. Results

5.2.1. Characterization of OfCPN1 cDNA sequence

The full-length sequence (1591 bp) of OfCPN1 contains a CDS of 1347 bp (including stop codon) encoding 488 aa (Fig. 5.1). The 5′-untranslated region (UTR) and 3′-UTR contain 232 bp and 72 bp, respectively. The encoded polypeptide contains a SP with 19 residues and important motifs, such as a Zn-binding region 1

signature ([81]PEFKYVGNMHGNEVLGRELLIKF[103]), Zn-binding region 2 signature ([220]HGGAVVANYPF[230]), peptidase-M14 Zn carboxypeptidase site, Zn-binding site 1, Zn- binding site 2 and N-glycosylation site. The predicted molecular mass of OfCPN1 was approximately 51 kDa, and the theoretical isoelectric point was 6.5.

```
          GAGCAGG AGAAAGAAGAGGACA AGAGGGGGAAAGAGC AAGCAGGGAGTAAGG  -172
AGTATGTTTGGTGTG TAAGGAGATAAACAT ACATCTGAAAGTGTT TTCCATAGTTGCAAG  -120
CAGTCTCTTCTCCCA GCCTCTCCTCTACTC CTGTGCTGCACTAAG TTGGGCTGAGGGACC   -60
ATGCAGCAGGGGTGG ACTCTTCCCTGGCTC GCTGCTCTCCTGCCA GGGCTGATGGGGCTC    60
 M  Q  Q  G  W  T  L  P  W  L  A  A  L  L  P  G  L  M  G  L       20
CTGGTGTGCGGCTCA GACTTCCAGCATCAC CGCTACGAGGACATG GTGCGAGCCCTGTTT   120
 L  V  C  G  S  D  F  Q  H  H  R  Y  E  D  M  V  R  A  L  F       40
GCAGTGCAGAGCGAA TGCCCCTACATCACG CGCATTTACAGCATC GGGCGCAGCGTGGAG   180
 A  V  Q  S  E  C  P  Y  I  T  R  I  Y  S  I  G  R  S  V  E       60
GGGCGCCACCTCTAC GTGCTGGAGTTCAGT GATAACCCGGGCATC CATGAAGCATTGGAG   240
 G  R  H  L  Y  V  L  E  F  S  D  N  P  G  I  H  E  A  L  E       80
CCGGAGTTCAAGTAC GTGGGCAACATGCAC GGCAACGAAGTGCTC GGCCGTGAGCTGCTC   300
 P  E  F  K  Y  V  G  N  M  H  G  N  E  V  L  G  R  E  L  L      100
ATTAAGTTCTCCCAG TTTCTCTGCGAGGAG TATCGGGCCGGAAAC CAGCGGATCATGAGA   360
 I  K  F  S  Q  F  L  C  E  E  Y  R  A  G  N  Q  R  I  M  R      120
CTGATCCATGACACG CGCATCCACATCCTG CCCTCCATGAACCCT GATGGCTACGAGGTG   420
 L  I  H  D  T  R  I  H  I  L  P  S  M  N  P  D  G  Y  E  V      140
GCTGCCAGACAGGGT CCAGAGTTCAACGGC TACCTGGTGGGTCGA GGGAACTCCAGAGAA   480
 A  A  R  Q  G  P  E  F  N  G  Y  L  V  G  R  G  N  S  R  E      160
ATTGATCTGAACCGG AACTTTCCAGACCTG AACGCACTCATGTAC TATTACGAGAAGACC   540
 I  D  L  N  R  N  F  P  D  L  N  A  L  M  Y  Y  Y  E  K  T      180
AACGGGCGAAACCAC CACCTACCCCTGCCG GACAACTGGGAGCAA CAGGTTGAACCAGAG   600
 N  G  R  N  H  H  L  P  L  P  D  N  W  E  Q  Q  V  E  P  E      200
ACCTTGGCGGTCATA AAATGGATGCAAAAC TACAATTTTATCCTG TCAGCCAACCTCCAT   660
 T  L  A  V  I  K  W  M  Q  N  Y  N  F  I  L  S  A  N  L  H      220
GGTGGAGCTGTGGTA GCCAATTACCCCTTC GACAAGTCGAGAGAT CCTCGCATTCGAGGG   720
 G  G  A  V  V  A  N  Y  P  F  D  K  S  R  D  P  R  I  R  G      240
AGGACCACGTATGCA GCCACTCCGGATGAC AAAATCTTCAGAAAG TTGGCGAGGACCTAC   780
 R  T  T  Y  A  A  T  P  D  D  K  I  F  R  K  L  A  R  T  Y      260
TCGTACGCTCACAGC TGGATGCACAAGGGT TGGAACTGTGGGGAC TTCTTTGATGAGGGG   840
 S  Y  A  H  S  W  M  H  K  G  W  N  C  G  D  F  F  D  E  G      280
ATCACCAACGGGGCC AGCTGGTACTCTCTT TCCAAAGGCATGCAG GACTTCAACTACCTG   900
 I  T  N  G  A  S  W  Y  S  L  S  K  G  M  Q  D  F  N  Y  L      300
TACAGCAACTGTTTT GAGATCACCCTAGAG CTGAGCTGTGATAAG TTTCCTCCAGCGTCA   960
 Y  S  N  C  F  E  I  T  L  E  L  S  C  D  K  F  P  P  A  S      320
GCACTGCCCAGAGAA TGGCTGGGCAACCGA GAAGCACTGGTTTCA TACCTTGAACAGGTG  1020
 A  L  P  R  E  W  L  G  N  R  E  A  L  V  S  Y  L  E  Q  V      340
CATCATGGGATAAAA GGCATGGTGTATGAT GAAAACAACAACCCT ATTGGCAACGCTGAG  1080
 H  H  G  I  K  G  M  V  Y  D  E  N  N  N  P  I  G  N  A  E      360
ATCTCAGTGTCTGGC ATCAACCACGATGTG ACCAGCGGAGTGGAT GGCGACTATTTCAGA  1140
 I  S  V  S  G  I  N  H  D  V  T  S  G  V  D  G  D  Y  F  R      380
CTTCTGTTACCAGGT ACCTACACTGTGACA GCATCTGCCCCAGGT TACGTTCCCTCCACC  1200
 L  L  L  P  G  T  Y  T  V  T  A  S  A  P  G  Y  V  P  S  T      400
AGCACTGTCACAGTG GGACCAGCTGAGGCC ATACAGCTTCATTTT TACTTGAAAACGGCA  1260
 S  T  V  T  V  G  P  A  E  A  I  Q  L  H  F  Y  L  K  T  A      420
CCAAAACAAAACCTG AAAGTGAAGCCCCAC AACGGCAAGAAGAAC CTCTCGTCTCCCAAG  1320
 P  K  Q  N  L  K  V  K  P  H  N  G  K  K  N  L  S  S  P  K      440
GCCCCGTTAAAACTC GGCCCTAGATGAGAC ATGGACAATAATTTA GACAACACACAAACA  1380
 A  P  L  K  L  G  P  R  *                                       448
TATATACACACACAC TGAAACTCCCACATA AAACATACG                        1419
```

Figure. 5.1. Nucleotide and deduced amino acid sequences of OfCPN1. The nucleotide sequence (upper) and the deduced amino acid sequence (lower) are numbered. The start codon (ATG) and stop codon (TGA) are boxed. The predicted SP is in bold type. The Zn-binding region 1 signature is boxed, and the Zn-binding region 2 signature is in bold and italic type. The peptidase-M14 Zn carboxypeptidase site is italicized and underlined. Zn-binding site 1 is indicated in grey, and the second Zn-binding site is marked with a hexagon. The N-glycosylation site is circled.

5.2.2. Pairwise and multiple alignment analysis of OfCPN1

Pairwise sequence analysis revealed that OfCPN1 has the highest identity (96.9 %) with European sea bass (*Dicentrarchus labrax*) CPN1 (CBN81456) (Table 5.1). The OfCPN1 identity is ranging 96.9 % to 84.9 % among the teleosts. Furthermore, OfCPN1 has 67.2 % identity with *Homo sapiens* CPN1 (NP_001299) where highest identity from mammals.

Table 5.1. Percent identity and similarity for the amino acid sequences of OfCPN1 from different organisms

	CPN1	% Identity												
		O. fasciatus	*D.labrax*	*O. niloticus*	*T. rubripes*	*O. latipes*	*S. salar*	*D. rerio*	*H. sapiens*	*X. tropicalis*	*M. musculus*	*B. taurus*	*R. norvegicus*	*P. alecto*
% Similarity	*Oplegnathus fasciatus*		**96.9**	93.1	90.9	89.3	87.9	84.9	67.2	66.4	65.6	65.2	64.5	64.3
	Dicentrarchus labrax	**97.8**		92.2	89.8	88.4	87.3	84.7	66.5	67.0	64.9	64.5	63.9	63.9
	Oreochromis niloticus	97.1	96.0		88.2	87.9	85.7	83.1	67.4	64.8	65.6	65.9	64.9	64.2
	Takifugu rubripes	94.4	93.3	93.8		84.9	84.8	82.7	64.8	65.6	64.4	63.8	64.0	62.7
	Oryzias latipes	93.5	92.9	93.3	91.3		84.2	81.6	65.6	66.5	64.6	64.3	63.8	62.7
	Salmo salar	93.1	92.9	91.1	90.6	90.6		83.8	65.8	66.3	64.2	64.3	64.0	63.0
	Danio rerio	90.4	90.2	89.1	88.7	89.1	91.3		66.0	66.9	64.6	65.0	63.4	62.7
	Homo sapiens	79.7	78.8	80.1	79.5	78.8	77.7	78.6		68.1	83.4	86.2	82.8	81.0
	Xenopus tropicalis	80.6	80.8	80.6	80.0	80.6	79.3	80.4	81.0		67.6	67.2	67.4	65.0
	Mus musculus	78.8	77.7	78.3	77.0	78.3	75.5	77.2	91.3	78.1		82.7	96.5	78.2
	Bos taurus	78.8	78.1	80.1	77.9	78.4	76.4	78.4	92.2	81.2	89.8		81.6	83.5
	Rattus norvegicus	77.5	76.4	77.7	76.6	77.5	75.5	76.6	91.3	77.9	98.0	88.7		77.8
	Pteropus alecto	76.0	75.8	76.9	75.8	75.8	75.6	76.4	88.3	77.9	86.0	89.4	86.2	

ClustalW multiple alignment tool was used to align the complete aa sequences of CPN1 (Fig. 5.2) orthologous. The CPN1 orthologous sequences were ranged from 447 to 462 aas. Zn-binding signature 1 and Zn-binding signature 2 were highly conserved selected CPN1 orthologous. Moreover, the two Zn-binding sites were

identical to those in the selected teleosts and other vertebrates. OfCPN1 PM14 Zn carboxypeptidase site has displayed little variation to the other orthologous.

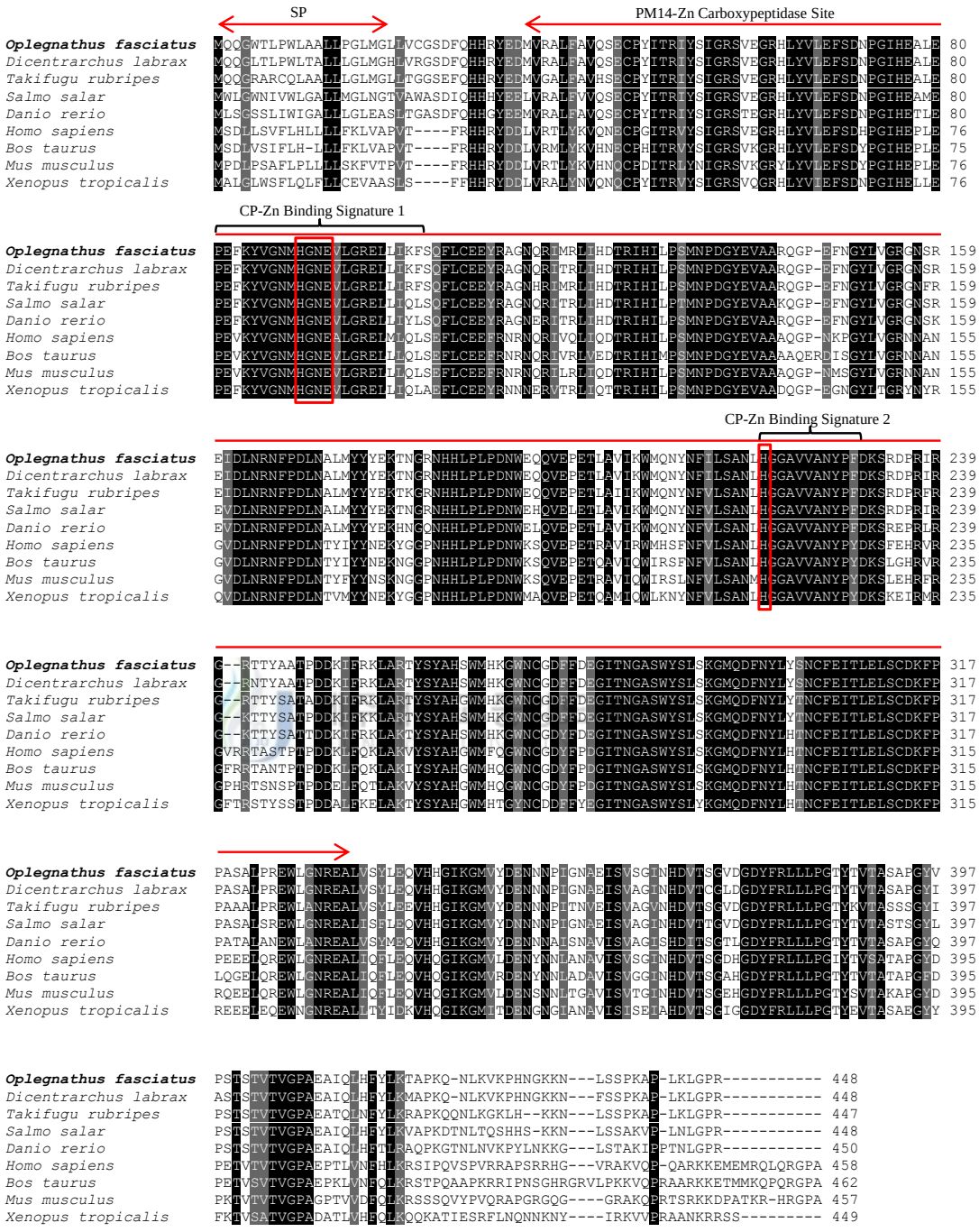


Figure. 5.2. Multiple sequence alignment of OfCPN1 with other known CPN1 AA sequences using ClustalW (v2.1). Identical residues are denoted by black shading. Similar residues are denoted by gray shading. Missing amino acids are indicated by dashes. The signal peptide (SP) and conserved PM14 Zn carboxypeptidase sites are marked by red color double arrows. The carboxypeptidase Zn binding signatures are denoted by brackets. The conserved, active Zn-binding sites are boxed within the Zn-binding signatures.

5.2.3. Molecular modeling

To determine the tertiary structure of OfCPN1, the I-TASSER ab-initio protein prediction algorithm was used with the corresponding top ten previously analyzed template crystal structures. The constructed tertiary structure of OfCPN1 contained 9 α-helices, 16 β-sheets, and 26 coils (Fig. 5.3). In a putative protein molecule, the sulfate ion was located in the center of the active site, near the site that is normally occupied by the catalytic zinc ion. The active site included interactions with side chains that are important for catalysis and for substrate recognition and fixation.

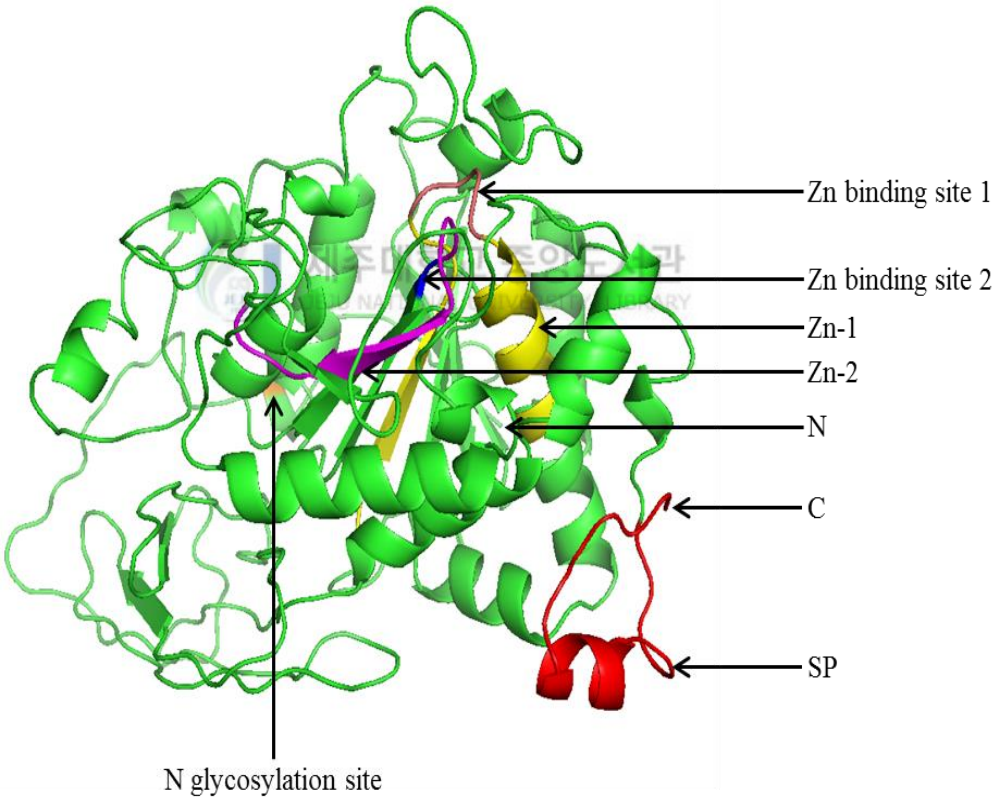


Figure. 5.3. Comparative tertiary structure of OfCPN1 putative protein molecule. The conserved SP is shown in red. The carboxypeptidase Zn-binding signature 1 (Zn-1) is marked in yellow, and the carboxypeptidase Zn-binding signature 2 (Zn-2) is colored in purple. The N-glycosylation site is indicated in orange, Zn-binding site 1 in Zn-binding signature 1 is shown in brown, and Zn-binding site 2 in Zn-binding signature 2 is shown in blue. The rest of the α-helices, β-sheets, and coils are in green.

5.2.4. Phylogenetic study

Phylogenetic analysis was carried out using the deduced aa sequences of different vertebrates to evaluate the evolutionary relationships of OfCPN1. The CPN1 homologues were cladded together with CPN1 homologues and OfCPN1 closely clustered with *D. labrax* (CBN81456) (Fig. 5.4) in the teleost groups. The results indicate that OfCPN1 evolved from the common ancestors of CPN1.

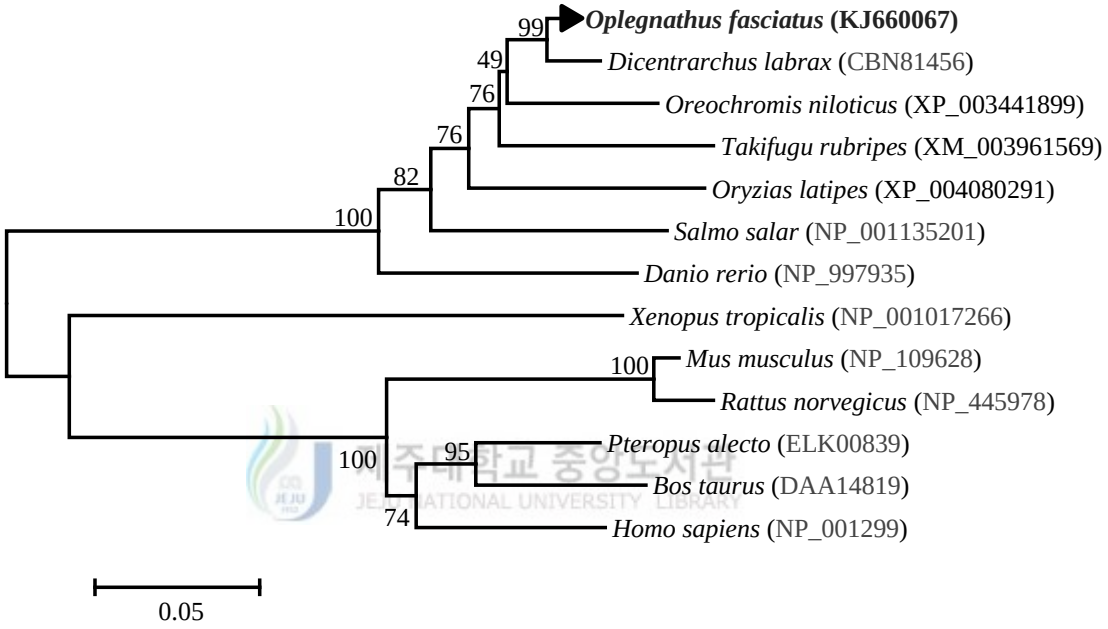


Figure. 5.4. Phylogenetic analysis of OfCPN1 with their known orthologues. The tree is based on the alignment of the full-length amino acid sequences. The tree was constructed using ClustalW (v2.1) and MEGA (v5.0). The numbers at the branches denote the bootstrap values for 5000 replicates. The GenBank accession numbers are shown within brackets next to each species.

5.2.5. Genomic structure analysis of OfCPN1

The genomic organization of CPN1 was analyzed and compared with the genomic structures of known counterparts, as shown in Fig. 5.5. The OfCPN1 genomic sequence contains 9 exons interrupted by 8 introns. Among the teleost CPN1 homologues, the 7 internal exons are conserved, but not the first and last exons in the CDS. The zebrafish *CPN1* genomic structure contains 7 exons at its 5′ end instead of

the single exon in OfCPN1. The coding region of OfCPN1 shows similarity with the other CPN1 homologues. The mammalian homologues contain a similar number of exons and introns, with six exons conserved among them. The exon lengths in chicken CPN1 are similar to the exon lengths (except last exon in CDS) in human and mouse CPN1.

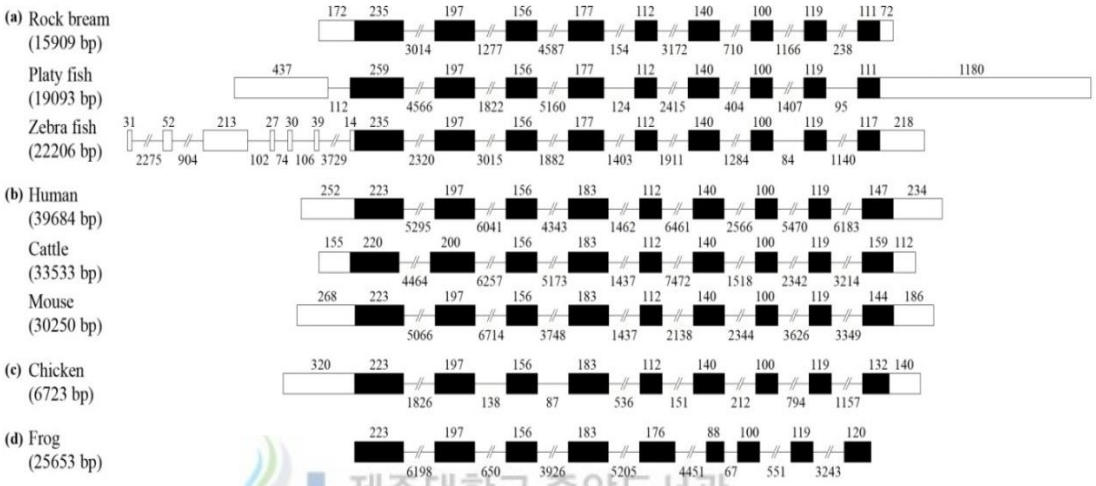


Figure. 5.5. Genomic structures of CPN1 (rock bream: **KJ660067**, platy fish: ENSXMAP00000018055, zebrafish: ENSDARP00000013414, human: ENSP00000359446, cattle: ENSBTAP00000032859, mouse: ENSMUSP00000026210, chicken: ENSGALP00000011468, frog: ENSXETP00000008089) from various vertebrates. Schematic exon-intron structures of teleosts (a), mammals (b), aves (c) and amphibians (d). Coding regions are shown as dark shaded boxes. The 5′- and 3′- untranslated regions are denoted with empty boxes, and introns are indicated by lines. The sizes of the exons are shown above the exon boxes, and the intron sizes are shown below the intron lines.

5.2.6. Tissue-specific mRNA expression profile of OfCPN1

The qPCR results showed that OfCPN1 were constitutively expressed in all tissues but at different levels (Fig. 5.6). The expression in each tissue was normalized to Ofβ-actin expression. The normalized OfCPN1 expression value in blood cells was set at 1, and expression in order tissues was expressed relative to these values. OfCPN1 transcript expression was highest in the liver and marginal in blood cells.

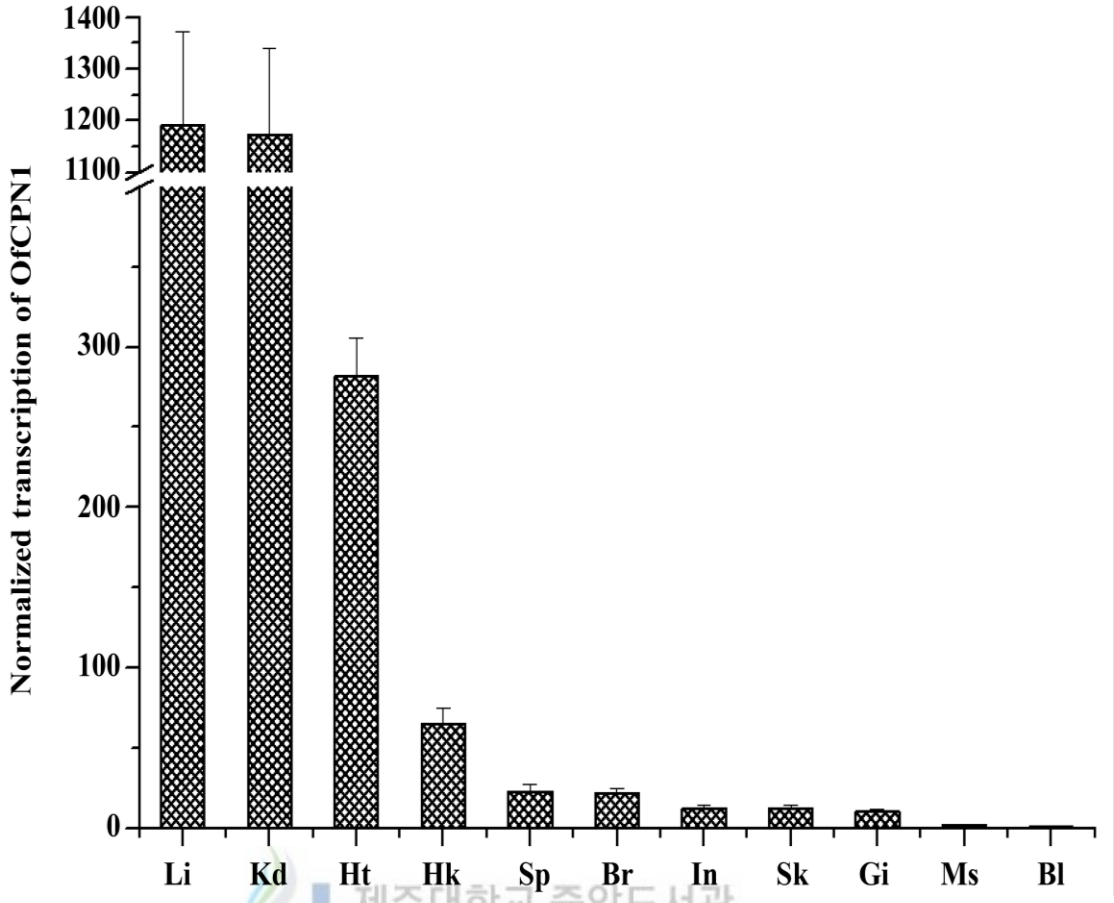


Figure. 5.6. Tissue-specific mRNA expression analysis of OfCPN1 by qPCR. The mRNA expression in the peripheral blood cells (Bl), gills (Gi), liver (Li), spleen (Sp), head kidney (Hk), kidney (Kd), skin (Sk), muscle (Ms), heart (Ht), brain (Br), and intestine (In) was normalized using Ofβ-actin as the reference gene. The error bars represent the standard deviation of triplicate samples. Significant differences in expression were calculated with respect to expression in blood ($P < 0.05$).

5.2.7. Immune responsive transcriptional regulation of OfCPN1

To evaluate the immune functions of OfCPN1, their mRNA levels in liver tissue were examined after challenge with *S. iniae*, *E. tarda*, RBIV, poly I:C, and LPS (Fig. 5.7). OfCPN1 mRNA expression showed significant elevation ($P < 0.05$) at various times p.i. in response to all immune regulators. Whereas, the most significant induction was detected in *E. tarda* challenge at 6 h p.i (fold ~ 4.9). However, RBIV induced a significant increase in OfCPN1 mRNA expression after the challenge, as compared to control treatment. Transcriptional level of OfCPN1 was significantly

higher in all the time points examined from 3 h to 48 h p.i, when compared to that of 0 h control. The expression of OfCPN1 after poly I:C challenge was highest at 48 h p.i. Moreover, it was also revealed that the OfCPN1 transcription at 48 h is statistically different ($P < 0.05$) from those of 3 h and 6 h.

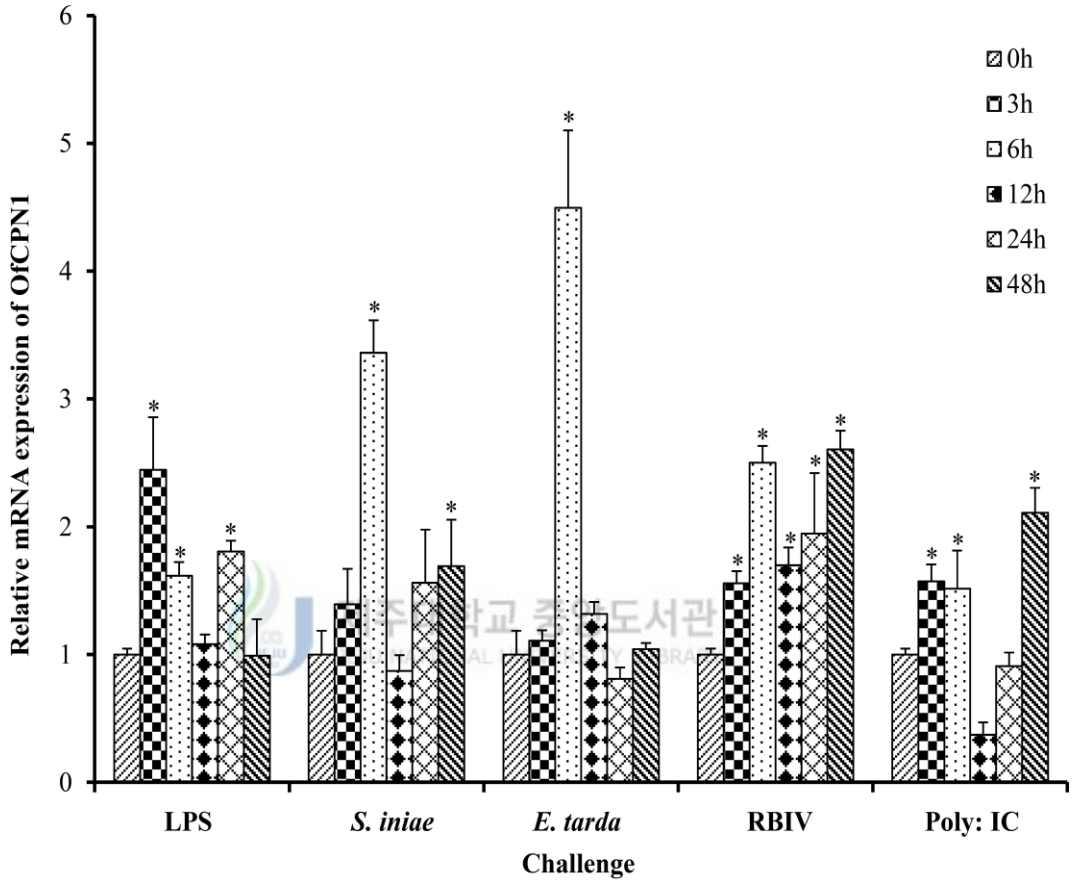


Figure. 5.7. The temporal mRNA expression of OfCPN1 in liver tissue from rock bream challenged with LPS, *Streptococcus iniae* (*S. iniae*), *Edwardsiella tarda* (*E. tarda*), rock bream iridovirus (RBIV), poly I:C, and PBS (control). The mRNA levels were normalized to Ofβ-actin mRNA levels. Gene expression in each control group (0 h) was set at 1, and expression in each challenge group was expressed relative to the expression in the respective PBS control group (expression > 1: upregulated; expression < 1: downregulated). mRNA levels at different time points were compared using Student's *t*-test. Data from three individuals were analyzed. Vertical bars represent the mean ± SD (n = 3). Significant differences ($P < 0.05$), relative to the 0 h values, are indicated with an asterisk.

5.3. Discussion

Immune homeostasis is maintained by different regulatory proteins for which other proteins cannot compensate. Among them, the CP family proteins have various biological functions in multicellular organisms of which several have been defined at the molecular level. The polypeptide encoded by OfCPN1 cDNA encoded a polypeptide of 448 aas, with a molecular weight of 51 kDa, similar to murine CPN1 (457 AAs, 51.845 kDa) (Matthews and Wetsel, 2001). OfCPN1 had the highest identity with *D. labrax*. Furthermore, OfCPN1 contained the characteristic Zn-binding signature motifs found in vertebrate CPs.

The active sites of CPN1 are highly conserved. Matthews *et al*. (2004) reported that mammalian CPs have various domains for Zn and substrate binding, peptide specificity, and catalytic activity. The Zn ion provides electrostatic stabilization for negatively charged intermediates, while other residues are important for substrate binding and catalysis. The sequence of the PM14 Zn carboxypeptidase site found in OfCPN1 was mostly similar to the sequences in teleosts and mammals. The M14 family includes Zn-binding CPs that hydrolyzes a single C-terminal aa and contains a recognition site for the free C-terminal carboxyl group, which is a key determinant of specificity. Given these findings, we suggest that the functions of OfCPN1 may similar from the functions of other vertebrate CPN1 proteins, which regulate innate immunity while acting as anaphylatoxin inactivator.

OfCPN1 is a member of the carboxypeptidase gene family; it was expected to have evolved from a common ancestral gene. This was affirmed by phylogenetic analysis. The fish CPN1 sequences were clustered together and clearly separated from the amphibian and mammalian CPN1 homologues sequences. The results indicate that the OfCPN1 gene duplications occurred before amphibians and mammals diverged.

Additionally, we found that OfCPN1 is most closely related to the CPN1 of *D. labrax*. The phylogenetic analysis in this study indicates a close evolutionary relationship among the fish CPN1 orthologues. From the evolution of the CPN1 genes, we can infer the evolutionary relationships between species to facilitate their classification.

Although the genomic sequences of many vertebrate CPN1 orthologues are present in the databases, only the genomic structures of mouse CPN1 (Matthews et al., 2004; Matthews and Wetsel, 2001) have been documented. This study describes the genomic structures of teleost CPN1 from rock bream and we compared the genomic structures of OfCPN1 with the genomic structures of CPN1 from seven other species. All followed the GT/AG exon/intron boundary splicing rule (Miranda-Vizuete and Spyrou, 2002). The CDS of teleost CPN1 comprises nine exons, as previously observed for mouse CPN1 (Matthews et al., 2004; Matthews and Wetsel, 2001). Aside from the first and last exons, the exons in OfCPN1 and other teleost CPN1 are conserved. Most of the exons within the CPN1 CDS are conserved among teleosts, birds, and mammals, confirming that the structures of CPN1 remained stable throughout their molecular evolution.

In the current study, qPCR showed constitutive expression of OfCPN1 in all the tissues examined. Moreover, it demonstrated a distinct tissue-specific transcriptional profile with varying amounts of mRNA in each tissue. OfCPN1 exhibited extremely strong expression in the liver (~1190-fold, relative to expression in blood), which are important organ associated with fish immunity.

Previously reported mouse CPN1 liver, has been reported to be the primary production site for CPN1. Furthermore, the OfCPN1 spatial mRNA expression profile was obtained in this study consistent with results of previous reports (Matthews et al.,

2004) (Matthews and Wetsel, 2001), indicating that the liver is the major site of OfCPN1 expression. More specifically, murine CPN is expressed in the liver, kidney, spleen, stomach, and intestine, but not in the brain tissue (Sato et al., 2000). In this study, OfCPN1 mRNA levels were high in the liver, heart, kidney, and head kidney, and negligible in the blood cells. The expression of OfCPN1 in all of the tissues examined, including non-immune tissues, indicates their importance in immune regulation and other cellular functions, which remain to be characterized.

OfCPN1 is an immune-related gene induced by different types of bacterial and viral immune stimulants. On the basis of the immune challenge results obtained from previous investigations, *S. iniae*, *E. tarda*, LPS, RBIV and poly I:C were used to induce inflammatory responses (Whang et al., 2011). Liver is the predominant site of carboxypeptidase gene expression and one of the central organs, which is prone to be attacked by many infectious microbes. Therefore, in order to examine the immune responses related to OfCPN1, their mRNA levels in liver tissue were determined post-immune challenges. The CPN is a crucial protein that can regulate the biological activities of complement anaphylatoxin (C3a, C4a, C5a) by cleaving the C-terminal arginine (Bokisch and Muller-Eberhard, 1970; Skidgel, 1995). The complement system, conserved in vertebrates and invertebrates, is a family of proteins that mediate potent biological responses in inflammation, immunity, and host defense (Matthews et al., 2004). Specifically, C3 plays a central role in the activation of the complement system, which promotes innate immunity with respect to bacterial infection (Matthews et al., 2004). Complement components such as C3a, C4a, and C5a have anaphylatoxin activity, including blood vessel dilation, increased vascular permeability, histamine release from mast cells, chemoattraction of eosinophils, chemotaxis of leukocytes, and induction of smooth muscle contraction (Matthews et

al., 2004). CPN1 regulates the anaphylatoxin activity of C3a, C4a, and C5a and protects tissues from damage caused by prolonged anaphylatoxin activity. In the immune challenge experiment, *S. iniae*, *E. tarda*, RBIV, poly I:C, and LPS triggered the early-phase expression of OfCPN1mRNA in liver tissue indicating a potential role for these proteins in the rock bream immune defense system. The expression patterns show that OfCPN1 are non-specific immune response genes in rock bream. Collectively, our findings indicate that OfCPN1 play a pivotal role in the immune responses to bacterial and viral infections in rock bream by acting as immune modulators. The results of this study provide molecular insights into CPN1 from teleost species and may help elucidate the functional mechanism of CPN1 in teleosts.

Conclusions

In the present study four major complement components have identified and characterized in terms of cDNA, gDNA levels, immune responsive transcriptional changes and functional assays within the one teleost species; rock bream, *Oplegnathus fasciatus*. Furthermore, employed experimental method provides complete and sound understanding of the immunological role of complement system as it summarized the mRNA expression of each gene upon diverse immune stimulants and pathogen challenges. Finally, the present approach in comprehensive study of four imperative complement components in rock bream fish *Oplegnathus fasciatus* will be important to understand its role in relation to survival of fish from challenging environment with abundant pathogenic threats and make significant insight into the immunological perspectives of the complement system in teleost.

References

Arlaud, G.J., Colomb, M.G., 2001. Complement: Classical Pathway, eLS. John Wiley & Sons, Ltd.

Arlaud, G.J., Gaboriaud, C., Garnier, G., Circolo, A., Thielens, N.M., Budayova-Spano, M., Fontecilla-Camps, J.C., Volanakis, J.E., 2002a. Structure, function and molecular genetics of human and murine C1r. Immunobiology 205, 365-382.

Arlaud, G.J., Gaboriaud, C., Thielens, N.M., Budayova-Spano, M., Rossi, V., Fontecilla-Camps, J.C., 2002b. Structural biology of the C1 complex of complement unveils the mechanisms of its activation and proteolytic activity. Molecular immunology 39, 383-394.

Arlaud, G.J., Gagnon, J., 1985. Identification of the peptide bond cleaved during activation of human C1r. FEBS letters 180, 234-238.

Arlaud, G.J., Rossi, V., Thielens, N.M., Gaboriaud, C., Bersch, B., Hernandez, J.F., 1998a. Structural and functional studies on C1r and C1s: new insights into the mechanisms involved in C1 activity and assembly. Immunobiology 199, 303-316.

Arlaud, G.J., Volanakis, J.E., Thielens, N.M., Narayana, S.V., Rossi, V., Xu, Y., 1998b. The atypical serine proteases of the complement system. Advances in immunology 69, 249-307.

Bathige, S.D., Whang, I., Umasuthan, N., Wickramaarachchi, W.D., Wan, Q., Lim, B.S., Park, M.A., Lee, J., 2013. Three complement component 1q genes from rock bream, Oplegnathus fasciatus: genome characterization and potential role in immune response against bacterial and viral infections. Fish & shellfish immunology 35, 1442-1454.

Batzoglou, S., Jaffe, D.B., Stanley, K., Butler, J., Gnerre, S., Mauceli, E., Berger, B., Mesirov, J.P., Lander, E.S., 2002. ARACHNE: a whole-genome shotgun assembler. Genome research 12, 177-189.

Baumann, H., Gauldie, J., 1994. The acute phase response. Immunology today 15, 74-80.

Beinrohr, L., Harmat, V., Dobo, J., Lorincz, Z., Gal, P., Zavodszky, P., 2007. C1 inhibitor serpin domain structure reveals the likely mechanism of heparin potentiation and conformational disease. The Journal of biological chemistry 282, 21100-21109.

Biesecker, G., Muller-Eberhard, H.J., 1980. The ninth component of human complement: purification and physicochemical characterization. Journal of immunology 124, 1291-1296.

Bokisch, V.A., Muller-Eberhard, H.J., 1970. Anaphylatoxin inactivator of human plasma: its isolation and characterization as a carboxypeptidase. The Journal of clinical investigation 49, 2427-2436.

Boshra, H., Li, J., Sunyer, J.O., 2006. Recent advances on the complement system of teleost fish. Fish & shellfish immunology 20, 239-262.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical biochemistry 72, 248-254.

Busby, T.F., Ingham, K.C., 1990. NH2-terminal calcium-binding domain of human complement C1s- mediates the interaction of C1r- with C1q. Biochemistry 29, 4613-4618.

Cammack, R., Atwood, T., Peter, C., Parish, H., Smith, A., Vella, F and Stirling, J.,, 2006. Oxford Dictionary of Biochemistry and Molecular Biology,. Oxford University Press, , 1586, Oxford, United Kingdom.

Carroll, S., Georgiou, G., 2013. Antibody-mediates inhibition of human C1s and the classical complement pathway. Immunobiology 218, 1041-1048.

Circolo, A., Garnier, G., Volanakis, J.E., 2003. A novel murine complement-related gene encoding a C1r-like serum protein. Molecular immunology 39, 899-906.

Davis, D.A., Singer, K.E., De La Luz Sierra, M., Narazaki, M., Yang, F., Fales, H.M., Yarchoan, R., Tosato, G., 2005. Identification of carboxypeptidase N as an enzyme responsible for C-terminal cleavage of stromal cell-derived factor-1alpha in the circulation. Blood 105, 4561-4568.

de Castro, E., Sigrist, C.J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P.S., Gasteiger, E., Bairoch, A., Hulo, N., 2006. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic acids research 34, W362-365.

Droege, M., Hill, B., 2008. The Genome Sequencer FLX System--longer reads, more applications, straight forward bioinformatics and more complete data sets. Journal of biotechnology 136, 3-10.

Erdos, E.G., Sloane, E.M., 1962. An enzyme in human blood plasma that inactivates bradykinin and kallidins. Biochemical pharmacology 11, 585-592.

Fearon, D.T., Locksley, R.M., 1996. The instructive role of innate immunity in the acquired immune response. Science 272, 50-53.

Gaboriaud, C., Rossi, V., Bally, I., Arlaud, G.J., Fontecilla-Camps, J.C., 2000. Crystal structure of the catalytic domain of human complement c1s: a serine protease with a handle. The EMBO journal 19, 1755-1765.

Gal, P., Barna, L., Kocsis, A., Zavodszky, P., 2007. Serine proteases of the classical and lectin pathways: similarities and differences. Immunobiology 212, 267-277.

Gal, P., Dobo, J., Zavodszky, P., Sim, R.B., 2009. Early complement proteases: C1r, C1s and MASPs. A structural insight into activation and functions. Molecular immunology 46, 2745-2752.

Garnier, G., Circolo, A., Xu, Y., Volanakis, J.E., 2003. Complement C1r and C1s genes are duplicated in the mouse: differential expression generates alternative isomorphs in the liver and in the male reproductive system. The Biochemical journal 371, 631-640.

Harboe, M., Ulvund, G., Vien, L., Fung, M., Mollnes, T.E., 2004. The quantitative role of alternative pathway amplification in classical pathway induced terminal complement activation. Clinical and experimental immunology 138, 439-446.

He, S., Yang, J.C., Tsang, S., Sim, R.B., Whaley, K., 1997. Role of the distal hinge region of C1-inhibitor in the regulation of C1s activity. FEBS letters 412, 506-510.

Holland, M.C., Lambris, J.D., 2002. The complement system in teleosts. Fish & shellfish immunology 12, 399-420.

Hsu, M.H., Ember, J.A., Wang, M., Prossnitz, E.R., Hugli, T.E., Ye, R.D., 1997. Cloning and functional characterization of the mouse C3a anaphylatoxin receptor gene. Immunogenetics 47, 64-72.

Jing, H., Babu, Y.S., Moore, D., Kilpatrick, J.M., Liu, X.Y., Volanakis, J.E., Narayana, S.V., 1998. Structures of native and complexed complement factor D: implications of the atypical His57 conformation and self-inhibitory loop in the regulation of specific serine protease activity. Journal of molecular biology 282, 1061-1081.

Journet, A., Tosi, M., 1986. Cloning and sequencing of full-length cDNA encoding the precursor of human complement component C1r. The Biochemical journal 240, 783-787.

Kardos, J., Gal, P., Szilagyi, L., Thielens, N.M., Szilagyi, K., Lorincz, Z., Kulcsar, P., Graf, L., Arlaud, G.J., Zavodszky, P., 2001. The role of the individual domains in the structure and function of the catalytic region of a modular serine protease, C1r. Journal of immunology 167, 5202-5208.

Kardos, J., Harmat, V., Pallo, A., Barabas, O., Szilagyi, K., Graf, L., Naray-Szabo, G., Goto, Y., Zavodszky, P., Gal, P., 2008. Revisiting the mechanism of the autoactivation of the complement protease C1r in the C1 complex: structure of the active catalytic region of C1r. Molecular immunology 45, 1752-1760.

Keil, C., Maskos, K., Than, M., Hoopes, J.T., Huber, R., Tan, F., Deddish, P.A., Erdos, E.G., Skidgel, R.A., Bode, W., 2007. Crystal structure of the human carboxypeptidase N (kininase I) catalytic domain. Journal of molecular biology 366, 504-516.

Kong, H.J., Hong, G.E., Nam, B.H., Kim, Y.O., Kim, W.J., Lee, S.J., Lee, N.S., Do, J.W., Cho, H.K., Cheong, J., Lee, C.H., Kim, K.K., 2009. An immune responsive complement factor D/adipsin and kallikrein-like serine protease (PoDAK) from the olive flounder Paralichthys olivaceus. Fish & shellfish immunology 27, 486-492.

Korotaevskiy, A.A., Hanin, L.G., Khanin, M.A., 2009. Non-linear dynamics of the complement system activation. Mathematical biosciences 222, 127-143.

Kusumoto, H., Hirosawa, S., Salier, J.P., Hagen, F.S., Kurachi, K., 1988. Human genes for complement components C1r and C1s in a close tail-to-tail arrangement. Proceedings of the National Academy of Sciences of the United States of America 85, 7307-7311.

Lacroix, M., Ebel, C., Kardos, J., Dobo, J., Gal, P., Zavodszky, P., Arlaud, G.J., Thielens, N.M., 2001. Assembly and enzymatic properties of the catalytic domain of human complement protease C1r. The Journal of biological chemistry 276, 36233-36240.

Lesavre, P.H., Muller-Eberhard, H.J., 1978. Mechanism of action of factor D of the alternative complement pathway. The Journal of experimental medicine 148, 1498-1509.

Letunic, I., Doerks, T., Bork, P., 2009. SMART 6: recent updates and new developments. Nucleic acids research 37, D229-232.

Levin, Y., Skidgel, R.A., Erdos, E.G., 1982. Isolation and characterization of the subunits of human plasma carboxypeptidase N (kininase i). Proceedings of the National Academy of Sciences of the United States of America 79, 4618-4622.

Li, H., Sun, Z.P., Li, Q., Jiang, Y.L., 2011. [Characterization of an iridovirus detected in rock bream (Oplegnathus fasciatus; Temminck and Schlegel)]. Bing du xue bao = Chinese journal of virology / [bian ji, Bing du xue bao bian ji wei yuan hui] 27, 158-164.

Lichtenfels, A.J., Lorenzi-Filho, G., Guimaraes, E.T., Macchione, M., Saldiva, P.H., 1996. Effects of water pollution on the gill apparatus of fish. Journal of comparative pathology 115, 47-60.

Lipton, D.W., Kim, D.H., , 2010. Accounting for economic risk and uncertainty in offshore aquaculture: a case study of Korean rock bream [Oplegnathus fasciats] production. Fish. Res. Agency.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402-408.

Mackinnon, C.M., Carter, P.E., Smyth, S.J., Dunbar, B., Fothergill, J.E., 1987. Molecular cloning of cDNA for human complement component C1s. The complete amino acid sequence. European journal of biochemistry / FEBS 169, 547-553.

Magnadottir, B., 2006. Innate immunity of fish (overview). Fish & shellfish immunology 20, 137-151.

Malhotra, R., Lu, J., Holmskov, U., Sim, R.B., 1994. Collectins, collectin receptors and the lectin pathway of complement activation. Clinical and experimental immunology 97 Suppl 2, 4-9.

Mathews, K.P., Curd, J.G., Hugli, T.E., 1986. Decreased synthesis of serum carboxypeptidase N (SCPN) in familial SCPN deficiency. Journal of clinical immunology 6, 87-91.

Matsushita, M., Fujita, T., 1992. Activation of the classical complement pathway by mannose-binding protein in association with a novel C1s-like serine protease. The Journal of experimental medicine 176, 1497-1502.

Matthews, K.W., Mueller-Ortiz, S.L., Wetsel, R.A., 2004. Carboxypeptidase N: a pleiotropic regulator of inflammation. Mol Immunol 40, 785-793.

Matthews, K.W., Wetsel, R.A., 2001. Characterization of mouse carboxypeptidase N small active subunit gene structure. Journal of immunology 166, 6196-6202.

Miranda-Vizuete, A., Spyrou, G., 2002. Genomic organization and identification of a novel alternative splicing variant of mouse mitochondrial thioredoxin reductase (TrxR2) gene. Molecules and cells 13, 488-492.

Muller-Eberhard, H.J., Gotze, O., 1972. C3 proactivator convertase and its mode of action. The Journal of experimental medicine 135, 1003-1008.

Nakao, M., Osaka, K., Kato, Y., Fujiki, K., Yano, T., 2001. Molecular cloning of the complement (C1r/C1s/MASP2-like serine proteases from the common carp (Cyprinus carpio). Immunogenetics 52, 255-263.

Nakao, M., Tsujikura, M., Ichiki, S., Vo, T.K., Somamoto, T., 2011. The complement system in teleost fish: progress of post-homolog-hunting researches. Developmental and comparative immunology 35, 1296-1308.

Neurath, H., 1986. The versatility of proteolytic enzymes. Journal of cellular biochemistry 32, 35-49.

Park Ii, S., 2009. Disease control in Korean aquaculture. Fish Pathology.

Pejler, G., Knight, S.D., Henningsson, F., Wernersson, S., 2009. Novel insights into the biological function of mast cell carboxypeptidase A. Trends in immunology 30, 401-408.

Plummer, T.H., Jr., Hurwitz, M.Y., 1978. Human plasma carboxypeptidase N. Isolation and characterization. The Journal of biological chemistry 253, 3907-3912.

Quiniou, S.M., Katagiri, T., Miller, N.W., Wilson, M., Wolters, W.R., Waldbieser, G.C., 2003. Construction and characterization of a BAC library from a gynogenetic channel catfish Ictalurus punctatus. Genetics, selection, evolution : GSE 35, 673-683.

Rabs, U., Martin, H., Hitschold, T., Golan, M.D., Heinz, H.P., Loos, M., 1986. Isolation and characterization of macrophage-derived C1q and its similarities to serum C1q. European journal of immunology 16, 1183-1186.

Reid, K.B., Colomb, M.G., Loos, M., 1998. Complement component C1 and the collectins: parallels between routes of acquired and innate immunity. Immunology today 19, 56-59.

Remington, S.J., Breddam, K., 1994. Carboxypeptidases C and D. Methods in enzymology 244, 231-248.

Rosen, B.S., Cook, K.S., Yaglom, J., Groves, D.L., Volanakis, J.E., Damm, D., White, T., Spiegelman, B.M., 1989. Adipsin and complement factor D activity: an immune-related defect in obesity. Science 244, 1483-1487.

Sakai, H., Nakashima, S., Yoshimura, S., Nishimura, Y., Sakai, N., Nozawa, Y., 1998. Molecular cloning of a cDNA encoding a serine protease homologous to complement C1s precursor from rat C6 glial cells and its expression during glial differentiation. Gene 209, 87-94.

Sato, T., Miwa, T., Akatsu, H., Matsukawa, N., Obata, K., Okada, N., Campbell, W., Okada, H., 2000. Pro-carboxypeptidase R is an acute phase protein in the mouse, whereas carboxypeptidase N is not. Journal of immunology 165, 1053-1058.

Skidgel, R.A., 1995. Human carboxypeptidase N: lysine carboxypeptidase. Methods in enzymology 248, 653-663.

Skidgel, R.A., Erdos, E.G., 2007. Structure and function of human plasma carboxypeptidase N, the anaphylatoxin inactivator. International immunopharmacology 7, 1888-1899.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology and evolution 28, 2731-2739.

Thielens, N.M., Bersch, B., Hernandez, J.F., Arlaud, G.J., 1999. Structure and functions of the interaction domains of C1r and C1s: keystones of the architecture of the C1 complex. Immunopharmacology 42, 3-13.

Tiron, A., Vasilescu, C., 2008. [Role of the spleen in immunity. Immunologic consequences of splenectomy]. Chirurgia 103, 255-263.

Umasuthan, N., Revathy, K.S., Bathige, S.D., Lim, B.S., Park, M.A., Whang, I., Lee, J., 2013. A manganese superoxide dismutase with potent antioxidant activity identified from Oplegnathus fasciatus: genomic structure and transcriptional characterization. Fish & shellfish immunology 34, 23-37.

Villiers, C.L., Arlaud, G.J., Colomb, M.G., 1985. Domain structure and associated functions of subcomponents C1r and C1s of the first component of human complement. Proceedings of the National Academy of Sciences of the United States of America 82, 4477-4481.

Volanakis, J.E., Narayana, S.V., 1996. Complement factor D, a novel serine protease. Protein science : a publication of the Protein Society 5, 553-564.

Walport, M.J., 2001. Complement. First of two parts. The New England journal of medicine 344, 1058-1066.

Wang, T., Secombes, C.J., 2003. Complete sequencing and expression of three complement components, C1r, C4 and C1 inhibitor, of the classical activation pathway of the complement system in rainbow trout Oncorhynchus mykiss. Immunogenetics 55, 615-628.

Wetsel, R.A., Kildsgaard, J., Haviland, D.L., 2000. Complement anaphylatoxins (C3a, C4a, C5a) and their receptors (C3aR, C5aR/CD88) as therapeutic targets in inflammation. In Therapeutic Interventions in the Complement System. Humana Press, Totawa.

Whang, I., Lee, Y., Lee, S., Oh, M.J., Jung, S.J., Choi, C.Y., Lee, W.S., Kim, H.S., Kim, S.J., Lee, J., 2011. Characterization and expression analysis of a goose-type lysozyme from the rock bream Oplegnathus fasciatus, and antimicrobial activity of its recombinant protein. Fish & shellfish immunology 30, 532-542.

White, R.T., Damm, D., Hancock, N., Rosen, B.S., Lowell, B.B., Usher, P., Flier, J.S., Spiegelman, B.M., 1992. Human adipsin is identical to complement factor D and is expressed at high levels in adipose tissue. The Journal of biological chemistry 267, 9210-9213.

Wimmers, K., Mekchay, S., Schellander, K., Ponsuksili, S., 2003. Molecular characterization of the pig C3 gene and its association with complement activity. Immunogenetics 54, 714-724.

Xu, Y., Narayana, S.V., Volanakis, J.E., 2001. Structural biology of the alternative pathway convertase. Immunological reviews 180, 123-135.

Yano, T., Nakao, M., 1994. Isolation of a carp complement protein homologous to mammalian factor D. Molecular immunology 31, 337-342.

Zhou, Z., Liu, H., Liu, S., Sun, F., Peatman, E., Kucuktas, H., Kaltenboeck, L., Feng, T., Zhang, H., Niu, D., Lu, J., Waldbieser, G., Liu, Z., 2012. Alternative complement pathway of channel catfish (Ictalurus punctatus): molecular characterization, mapping and expression analysis of factors Bf/C2 and Df. Fish & shellfish immunology 32, 186-195.