Available online at www.sciencedirect.com
ScienceDirect
Molecular Immunology 44 (2007) 3005–3023
Molecular Immunology
www.elsevier.com/locate/molimm

Short and long peptidoglycan recognition proteins (PGRPs) in zebrafish, with findings of multiple PGRP homologs in teleost fish

M.X. Chang, P. Nie*, L.L. Wei

State Key Laboratory of Freshwater Ecology and Biotechnology, and Laboratory of Fish Diseases, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei Province 430072, PR China

Received 1 November 2006; received in revised form 20 December 2006; accepted 20 December 2006
Available online 12 February 2007

Abstract

Peptidoglycan recognition protein (PGRP) specifically binds to peptidoglycan and is considered to be one of the pattern recognition proteins in the innate immunity of insect and mammals. Using a database mining approach and RT-PCR, multiple peptidoglycan recognition protein (PGRP) like genes have been discovered in fish including zebrafish *Danio rerio*, Japanese pufferfish *TakiFugu rubripes* and spotted green pufferfish *Tetraodon nigroviridis*. They share the common features of those PGRPs in arthropod and mammals, by containing a conserved PGRP domain. Based on the predicted structures, the identified zebrafish PGRP homologs resemble short and long PGRP members in arthropod and mammals. The identified PGRP genes in *T. nigroviridis* and *TakiFugu rubripes* resemble the long PGRPs, and the short PGRP genes have not been found in *T. nigroviridis* and *TakiFugu rubripes* databases. Computer modelling of these molecules revealed the presence of three α-helices and five or six β-strands in all fish PGRPs reported in the present study. The long PGRP in teleost fish have multiple alternatively spliced forms, and some of the identified spliced variants, e.g., tnPGRP-L3 and tnPGRP-L4 (tn: *Tetraodon nigroviridis*), exhibited no characters present in the PGRP homologs domain. The coding regions of zfPGRP6 (zf: zebrafish), zfPGRP2-A, zfPGRP2-B and zfPGRP-L contain five exons and four introns; however, the other PGRP-like genes including zfPGRPSC1a, zfPGRPSC2, tnPGRP-L1-, tnPGRP-L2 and frPGRP-L (fr: *Takifugu rubripes*) contain four exons and three introns. In zebrafish, long and short PGRP genes identified are located in different chromosomes, and an unknown locus containing another long PGRP-like gene has also been found in zebrafish, demonstrating that multiple PGRP loci may be present in fish. In zebrafish, the constitutive expressions of zfPGRP-L, zfPGRP-6 and zfPGRP-SC during ontogeny from unfertilized eggs to larvae, in different organs of adult, and the inductive expression following stimulation by *Flavobacterium columnare*, were detected by real-time PCR, but the levels and patterns varied for different PGRP genes, implying that different short and long PGRPs may play different roles in innate immune response.

Keywords: Peptidoglycan recognition protein; Long PGRP; Short PGRP; Fish; Zebrafish; Japanese pufferfish; Spotted green pufferfish

1. Introduction

The immune system has traditionally been divided into innate and adaptive components. The innate immunity, which includes antimicrobial peptides, phagocytes, alternative complement pathway, etc., plays vital roles in primary defense against invading pathogens in both vertebrates and invertebrates. The strategy of innate immune response is to recognize highly conserved structures present in a large group of microorganisms but rare or absent in responding species. These structures are referred to as pathogen-associated molecular pattern, and the receptors in innate immune system, which have evolved to recognize these structures, are called pattern recognition receptors (PRRs; Medzhitov and Janeway, 1997).

In insects and mammals, a number of proteins which are pattern recognition receptors have been described. The PRRs identified in mammals include C-type lectin, proteins with leucine regions, scavenger receptors, pentraxins, lipid transferases, integrins and complement control proteins (Medzhitov and Janeway, 1997). In *Anopheles* and *Drosophila*, the PRRs identified can be divided into six gene families: peptidoglycan recognition proteins (PGRPs), thioester-containing proteins (TEPs), gram-negative binding proteins (GNBP), the multidomain scavenger receptors (SCR), C-type lectins (CTL) and

* Corresponding author. Tel.: +86 27 68780736; fax: +86 27 68780123.
E-mail address: pinnie@ihb.ac.cn (P. Nie).

0161-5890/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.molimm.2006.12.029

galectins (GALE) (Christophides et al., 2002). These PRRs play important roles in facilitating phagocytosis and signal transduction pathways leading to the synthesis of anti-pathogen effectors.

PGRPs which are evolutionarily highly conserved pattern recognition receptors can bind specifically to peptidoglycan moieties. The first member of the PGRP family was discovered in silkworm *Bombyx mori* as a protein that initiates the activation of prophenol-oxidase cascade in the presence of peptidoglycan (Yoshida et al., 1996). Subsequently, its homologous proteins have been identified in other lepidopteran insects (Kang et al., 1998) and in diptera *Drosophila* (Werner et al., 2000). PGRP homologs have also been reported in mammals such as *Homo sapiens* (Liu et al., 2001), *Mus musculus* (Liu et al., 2000), *Rattus norvegicus*, *Camelus dromedarius* and *Sus scrofa* (Sang et al., 2005). The recent completion of the Japanese pufferfish *Takifugu rubripes* and zebrafish *Danio rerio* genomes has enabled the identification of several PGRPs in fish. However, in spite of the sequence information from *T. rubripes* and *D. rerio* genomes, the research on fish counterparts of mammalian PGRPs has not yet been reported. Taking the advantage of expressed sequence tag (EST) databases and the completed fish genomes including *D. rerio*, *T. rubripes* and *Tetraodon nigroviridis*, PGRP genes were investigated in teleost fish in the present study. In addition, expression patterns of peptidoglycan recognition protein homologs were analyzed in oocytes, fertilized eggs, and larval stages from 1 to 60 days post-hatching and in different adult organs of zebrafish. The expression level was also compared between healthy fish and fish infected with *Flavobacterium columnare*, a pathogen causing columnaris disease of many species of fish.

2. Materials and methods

2.1. Database mining of fish peptidoglycan recognition protein sequences

The BLAST server at the National Center for Biotechnology Information (NCBI) was used to check the *D. rerio*, *Takifugu rubripes* and *Tetraodon nigroviridis* EST and genome sequences for the possible presence of PGRP domain encoding sequences. To predict intron–exon boundaries of the teleost fish genes, the GENSCAN web server at the Massachusetts Institute of Technology (http://www.genes.mit.edu/burgelab/) or the GENSCAN predictions of the Sanger Ensembl were used. The predicted sequences were adjusted manually based on comparisons with the homologous human genes.

2.2. PCR cloning

Five larvae of 60-day-old zebrafish and the liver of *T. nigroviridis* were homogenized in Trizol Reagent (Invitrogen, USA), with total RNA extracted according to the manufacturer's instruction. The cDNA was synthesized using a Clontech SMART PCR cDNA Synthesis Kit (Clontech). To confirm coding sequence of peptidoglycan recognition protein 2, PCR was performed with primers zfPGRP2-F and zfPGRP2-R (Table 1). Similarly, the coding sequence of peptidoglycan recognition protein 6 was obtained by PCR using primers zfPGRP6-F and zfPGRP6-R, using primers zfPGRP-SCF and zfPGRP-SCR for peptidoglycan recognition protein SC, and using primers tnPGLF and tnPGLR for *T. nigroviridis* PGRP-L.

Table 1
Primer sequences

Names	Sequences	Annealing temperature (°C)
UPM	CTAATACGACTCACTATAGGGC	
zfPGRP-LF	TCACTCACGGTTTTATG	52
zfPGRP-LR	TTTGATTCCTTTTGCTA	
zfPGRP-LRout	TTCGGATGAGTTTCAGTTTCGTTG	64/62
zfPGRP-LRin	CCGTGAGTGACATTGTGTGTTTGG	
zfPGRP-LFout	AAGGGGCTCACACTAAAGGACGCA	64/62
zfPGRP-LFin	AGGAGGATTTCACCATTCTCGGAC	
zfPGRP6-F	GTGGTGCTGACTTTGGAT	52
zfPGRP6-R	GTTGCTCTGCTGGTGGTA	
zfPGRP2-F	CAACCTCAATCTCTCTCA	49
zfPGRP2-R	TACCATCAATCTTTCCTC	
zfPGRP-SCF	GTCATGTTGGAGCATACAGGG	58
zfPGRP-SCR	GTCAAACACATTCACACAGTTC	
tnPGLF	TCCATTTCTTGGGTCCTTCA	55
tnPGLR	TGACACGCCGTATCCTAAACT	
ExzfPGRP-LF	TGGCAATACCCAGCAAACCCTGTC	58
ExzfPGRP-LR	GTCCTTTAGTGTGAGCCCCTTGT	
ExzfPGSC-F	ACATAAACGCAGATACAGTGAGTC	56
ExzfPGSC-R	GCTCCTACAATCCCCCATCCTC	
ExzfPGRP6-F	AACTCTATTGGCTACGGCG	54
ExzfPGRP6-R	TGTTTCCAGGACACTCGGT	
GAPDH-F	GTAACTCCGCAGAAAAGCCAGAC	58
GAPDH-R	CAAAAGAAACTAACACACACACA	

According to the EST sequence (GenBank accession no. CF999082), the internal region of zfPGRP-L was amplified by primer pairs zfPGRP-LF and zfPGRP-LR. The full-length cDNA sequence of zfPGRP-L was amplified by random amplification of cDNA ends (RACE). The 5′ end region of the zfPGRP-L was amplified by nested PCR with two primer pairs of UPM/zfPGRP-LRout (first round PCR) and UPM/zfPGRP-LRin (second round). The resultant products were isolated using the Omega agarose purification Kit, and cloned into pMD-18 Vector (TaKaRa) by following manufacturer's instructions. Putative clones were screened by PCR using the above primers under the same cycle conditions, and the selected clones were sequenced using the dideoxy chain-termination method on an automatic DNA sequencer (ABI Applied Biosystems Model 377). The gene specific primers used for 3′ RACE PCR were UPM/zfPGRP-LFout (first round PCR) and UPM/zfPGRP-LFin (second round). The PCR program for RACE was as the followings: 1 cycle of 94 °C for 3 min, 5 cycles of 94 °C for 30 s, 66 °C for 30 s, and 72 °C for 120 s, 30 cycles of 94 °C for 30 s, 62 °C for 30 s, and 72 °C for 120 s, followed by 1 cycle of 72 °C for 7 min. The PCR products were cloned and sequenced as described above. All primers were shown in Table 1.