



Review

# Pathogen-Derived Carbohydrate Recognition in Molluscs Immune Defense

Weilin Wang<sup>1,2</sup>, Xiaorui Song<sup>1</sup>, Lingling Wang<sup>1</sup> and Linsheng Song<sup>1,2,3,\*</sup>

<sup>1</sup> Liaoning Key Laboratory of Marine Animal Immunology, Dalian Ocean University, Dalian 116023, China; wangweilin@dlou.edu.cn (W.W.); songxiaorui@dlou.edu.cn (X.S.); wanglingling@dlou.edu.cn (L.W.)

<sup>2</sup> Functional Laboratory of Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266235, China

<sup>3</sup> Liaoning Key Laboratory of Marine Animal Immunology and Disease Control, Dalian Ocean University, Dalian 116023, China

\* Correspondence: lshsong@dlou.edu.cn; Tel.: +86-411-8476-3173

Received: 23 January 2018; Accepted: 22 February 2018; Published: 3 March 2018

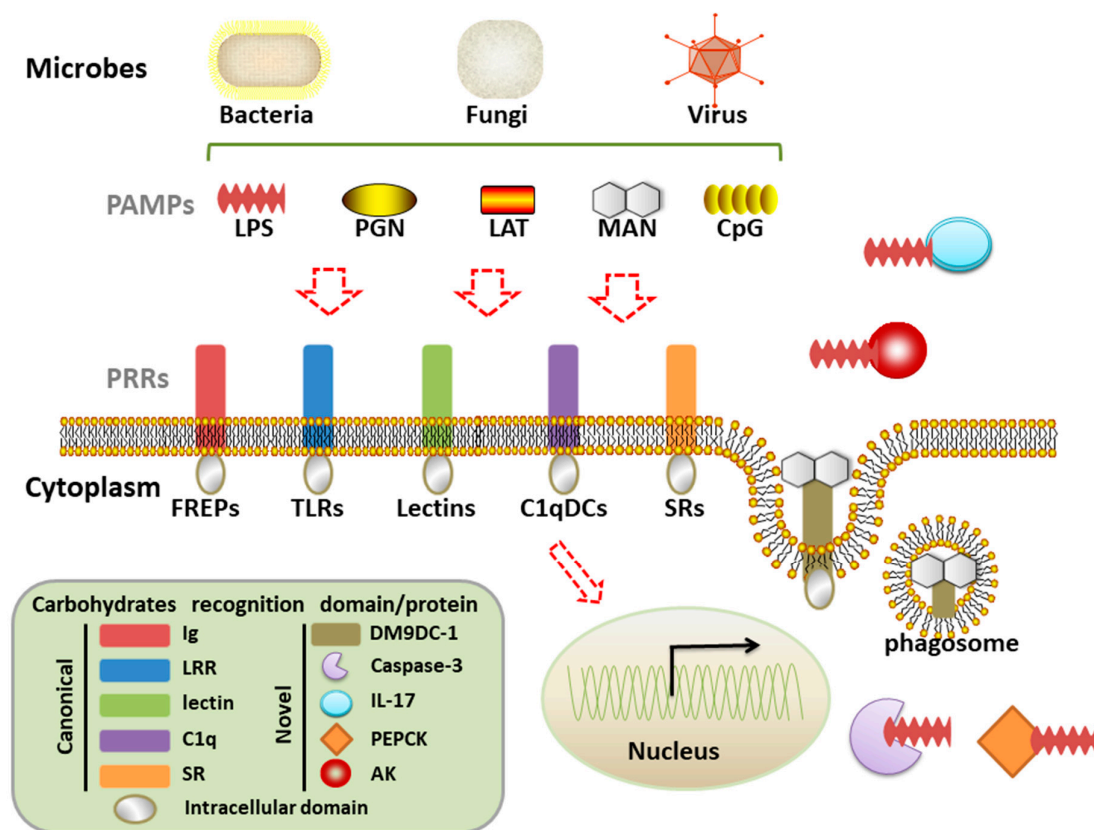
**Abstract:** Self-nonsel self discrimination is a common theme for all of the organisms in different evolutionary branches, which is also the most fundamental step for host immune protection. Plenty of pattern recognition receptors (PRRs) with great diversity have been identified from different organisms to recognize various pathogen-associated molecular patterns (PAMPs) in the last two decades, depicting a complicated scene of host-pathogen interaction. However, the detailed mechanism of the complicate PAMPs–PRRs interactions at the contacting interface between pathogens and hosts is still not well understood. All of the cells are coated by glycosylation complex and thick carbohydrates layer. The different polysaccharides in extracellular matrix of pathogen-host are important for nonself recognition of most organisms. Coincidentally, massive expansion of PRRs, majority of which contain recognition domains of Ig, leucine-rich repeat (LRR), C-type lectin (CTL), C1q and scavenger receptor (SR), have been annotated and identified in invertebrates by screening the available genomic sequence. The phylum Mollusca is one of the largest groups in the animal kingdom with abundant biodiversity providing plenty of solutions about pathogen recognition and immune protection, which might offer a suitable model to figure out the common rules of immune recognition mechanism. The present review summarizes the diverse PRRs and common elements of various PAMPs, especially focusing on the structural and functional characteristics of canonical carbohydrate recognition proteins and some novel proteins functioning in molluscan immune defense system, with the objective to provide new ideas about the immune recognition mechanisms.

**Keywords:** carbohydrate complex; pattern recognition receptor; molluscs; immune recognition; innate immunity

## 1. Introduction

All of the organisms, including prokaryote, plant, and animal are facing potentially life-threatening infections by a variety of microbial pathogens. The survival of the host depends on its ability to recognize infectious microbes and induce appropriate defense responses [1,2]. The innate immune recognition is a fundamental and core reaction to sense invaders and trigger the following immune protection. Since the discovery of Toll like receptors (TLRs) in 1990s, the mechanism of immune recognition between hosts and pathogens has been gradually uncovered, with the pattern recognition mechanism proposed in 2002 as a notable milestone [3]. So far, various pattern recognition receptors (PRRs) have been reported to induce immune defense in plants, invertebrates, and vertebrates by recognizing conserved structures of different pathogens, called pathogen-associated molecular patterns (PAMPs).

In the long-term evolution, the immune protection first starts with non-professional immune cells in plant, cnidarians, and nematode, and the epithelial cells are able to mediate innate immune responses [4]. The true cavity is developed from the endomesoderm of triploblastica and is filled with body fluid, which is concomitant with the presence of open circulatory system and specialized immunocytes (hemocytes) in ancient mollusc [5,6]. The molluscan immune system could be inferred as the most ancient prototype of highly specialized immune system of advanced vertebrates. Moreover, the phylum Mollusca is one of the largest groups in the animal kingdom with abundant biodiversity, which provides plenty of natural solutions about pathogens defense [7]. Accordingly, the study of molluscan immune recognition will be of great reference for understanding the common codex behind nonself recognition. In this concern, the present review introduces the common elements of various PAMPs and diverse PRRs, with the emphasis on the canonical and novel carbohydrate recognition proteins in molluscs (Figure 1).



**Figure 1.** Canonical and novel carbohydrate recognition proteins functioning in molluscan immune defense system. The canonical pattern recognition receptors (PRRs) are mainly composed of five dominant recognition domains, e.g., Ig domain, leucine-rich repeat (LRR) domain, lectin domain, C1q domain, and scavenger receptor (SR) domain. By recognizing (red dotted arrows) the conserved pathogen-associated molecular patterns (PAMPs) of bacteria, fungi and viruses, these PRRs could activate cellular or humoral immune defense and protect host from infection. Except for the canonical PRRs, some new proteins were identified with pathogen derived carbohydrate recognition activity in molluscs, including novel DM9 domain containing protein 1 (DM9DC-1) and some other well-known proteins whose PRR function was discovered recently. (Lipopolysaccharide, LPS; Peptidoglycan, PGN; Lipoteichoic acids, LTA; Mannan, MAN; Fibrinogen-related proteins, FREPs; Toll like receptors, TLRs; C1q domain containing proteins, C1qDCs; Scavenger receptors, SRs; Phosphoenolpyruvate carboxykinase, PEPCK; Arginine kinase, AK; black arrow, gene expression).

## 2. Carbohydrate as a Main Element of PAMPs in Microbes

There are four major classes of biological macromolecules that are required for all the living organisms, including carbohydrates, proteins, nucleic acids, and lipids. They are composed of different types of subunit (sugars, amino acids, nucleotides, and fatty acids) and exhibit different profile in component, structure, and distribution. For example, the polysaccharides in the extracellular matrix (ECM) of algae, plants, bacteria, fungi and animals are different, such as cellulose in plants, peptidoglycan (PGN) in bacteria, chitin and glucans in fungi, and sulfated polysaccharides in animals [8]. The four individual macromolecules or their complex constitute the main elements of PAMPs in diverse pathogens. For example, lipopolysaccharide (LPS), peptidoglycan (PGN), lipoteichoic acids (LTA), and cell-wall lipoproteins are the main bacterial PAMPs.  $\beta$ -glucan (GLU) and mannan (MAN), the typical components on cell walls of fungi, are the fungal PAMPs [9]. Ribonucleic acid or deoxyribonucleic acid, which are often mimicked by poly I:C or CpG, are considered as the main targets for the immune system to recognize virus. These PAMPs show a common feature in that their chemical bases are all derived from one of the four macromolecules or two complexes (Table 1).

**Table 1.** The main component of common pathogen-associated molecular patterns (PAMPs).

PAMPs	Pathogens	Carbohydrates	Proteins	Nucleic Acids	Lipids
LPS	G−	+			+
PGN	G+	+	+		
LTA	G+	+	+		
Flagellin	G+/G−	+			
Lipoprotein	G+/G−	+	+		
MAN	fungi	+			
Poly I:C	virus			+	
CpG	virus			+	

Lipopolysaccharide, LPS; peptidoglycan, PGN; mannan, MAN; lipoteichoic acid, LTA.

Carbohydrates (glycans), as the most stable organic macromolecule with complex structure, have long been known to play fundamental roles in antigen formation and innate immune recognition [10,11]. Glycosylation is essential for almost all of the organisms, including fungi, yeasts, plants, insects, fishes, birds, and mammals, and most of their secreted and cell-surface proteins are glycosylated [12]. Meanwhile, all of the cells are coated with thick layers of complex carbohydrates, known as the glycocalyx, in which the glycan components are present in many different glycoforms, such as glycoproteins, proteoglycans, and glycolipids [13]. These glyco-complexes either secreted or located on the cell surface are the main molecular components of pathogen-host interaction, and essential for both pathogen invading and host innate recognition [14]. For example, LPS is the common component on the cell surface of Gram-negative bacteria, which can be recognized by TLR, C-type lectin (CTL), and scavenger receptor of host innate immunity as PAMP to recognize the invaded bacteria [3]. The other carbohydrates as main elements of PAMPs in microbes are also the ligands of various PRRs of host immunity to recognize the pathogens.

## 3. Main Recognition Domains of PRRs in Various Organisms

With the long evolution, all organisms have developed an effective immune recognition system, in which several families of germline-encoded PRRs can detect the conserved molecular patterns of microbes. Recently, a picture of the evolutionarily conserved sensing system has emerged by comparing the structural homologies of these pathogen sensors across species [15]. Most PRR families with evolutionarily conserved structures execute function by forming common motif and binding sites. For example, toll like receptors (TLRs) in simple nematode, diverse molluscs, and complex mammals recognize various PAMPs through the conserved LRR domains.

In terms of structure, almost all of the PRRs contain specific domains classified into different types. Immunoglobulin (Ig) domain, leucine-rich repeat (LRR) domain, C-type lectin (CTL) domain, C1q domain, and scavenger receptor (SR) cysteine-rich domain are identified as the most prominent motifs, which are also rearranged with various combinations to build immune recognition rhyme [15]. For example, fibrinogen-related proteins (FREPs) are composed of tandem Ig domains, and TLRs are composed of varied number of LRR domains (Table 2). The recognition characters of each domain endow a PRR with functions to recognize and discriminate specific ligands.

**Table 2.** The main recognition domains of various pattern recognition receptors (PRRs).

PRRs	Ig	LRR	Lectin	C1q	SR
FREP	+				
TLR		+			
CTL			+		
C1qDC				+	
SR					+

Fibrinogen-related protein, FREP; Toll like receptor, TLR; C-type lectin receptor, CTL; C1q domain containing protein, C1qDC; scavenger receptor, SR.

Recently, species-specific expansion of recognition domains has been significantly characterized in plants and invertebrates by screening the available genomic sequence. There is a big family of LRR containing recognition proteins reported in plant *Arabidopsis* [16], and 187 FREPs with Ig domains annotated in sea anemone *Nematostella vectensis* [17]. In oyster *Crassostrea gigas*, there is significant expansion of C1q, CTL, and Ig recognition domains, and 321 C1qDCs, 266 CTLs, and 190 FREPs [17] are annotated in its genome. The massive increases of LRR (222 TLRs and 203 NLRs), SR (218 SRs) and CTL (283 CTLs) recognition domains are also reported in sea urchin *Strongylocentrotus purpuratus* [17]. These recognition domain containing proteins play an indispensable role in the first step of immune defense, and the research progresses have greatly expanded our understanding of pattern recognition mechanisms. Some of their detailed characteristics and functions in carbohydrate recognition are discussed below.

#### 4. Canonical Carbohydrates Recognition Protein in Molluscs

The sophisticated repertoire of PRRs in molluscs could be mainly classified into several families, depending on their recognition domains, such as Ig domain containing PRRs (FREPs, IgSF proteins), LRR domain containing PRRs (TLRs and LRRs), lectin domain containing PRRs (CTL, Galectin), C1q domain containing PRRs (C1qDC), SR domain containing PRRs (SRs), and other domain containing PRRs (PGRPs, GNBP) (Table 2). These molluscan PRRs are canonical protein to recognize carbohydrates-complex of invaders.

##### 4.1. Ig Domain Containing Pattern Recognition Receptors

Ig domain is an evolutionarily ancient domain, and it can tandemly build a functional protein or assemble complexes with other domains. An individual Ig domain is approximately 100 amino acids in length and forms an "Ig-fold" structure, consisting of two anti-parallel  $\beta$ -sheets packed face to face [18]. Four sets of Ig domains have been described as variable-like domains (V), constant-like domains (C1 and C2), and intermediate domains (I) [18]. So far, varied numbers of Ig superfamily (IgSF) have been identified in different phylogenetically organisms, including FREPs in molluscs [19,20], Dscams in arthropods, and Variable region-containing chitin-binding proteins (VCBPs) in echinoderm [21]. Though lacking Ig dependent adaptive immunity, molluscs might be endowed with enhanced innate immunity by the recognition, adhesion, and opsonic roles of these Ig domain containing proteins.

#### 4.1.1. Fibrinogen-Related Proteins (FREPs)

FREP is a kind of Ig and fibrinogen-like (FBG) domain containing protein with high levels of sequence diversity, and it is involved in innate defense responses of invertebrates [22]. FREPs have been recently identified in a variety of mollusc, such as oyster *C. gigas* [19,23], snail *Biomphalaria glabrata* [24,25], and marine opisthobranch *Aplysia californica* [26]. There are about 190 predicted FREP genes with more than 200 FBG domains and 70 FBG-encoding genes identified in the genome of *C. gigas* [23] and *Lottia gigantea* [27], respectively. The large amount of FREPs suggests that they have considerable functional significance in the molluscan immune system. In addition, just like Dscam molecules in arthropods, the FREPs reported in molluscs are of abundant sequential polymorphism in their Ig domains. In *B. glabrata*, a surprising diversity of FREP genes were generated through point mutation and recombination [28], and the finding of alternatively spliced forms of FREPs transcripts and retrosequences suggested that their diversification occurred at both genomic and transcriptional levels [24,29]. But, the detailed mechanisms of FREPs polymorphism still need further studies.

Molluscan FREPs have been reported to be involved in immune response as an important kind of PRRs, mainly with binding and opsonic activities [30]. The mRNA expression of molluscan FREPs can be induced by PAMP stimulations, and FREP proteins exhibit varied binding activities towards diverse microbes. The mRNA expression level of FREPs in mussel *M. galloprovincialis* was significantly up-regulated after LPS, LTA, poly I:C, and zymosan stimulations [31]. AiFREP-1 from scallop *Argopecten irradians* could agglutinate Gram-negative bacteria *E. coli* JM109, *V. anguillarum* and Gram-positive bacteria *M. luteus* in the presence of calcium ions [32]. AiFREP-2 could bind different PAMP ligands including LPS, PGN and GLU, and various microbes, including Gram-negative bacteria (*V. anguillarum*), Gram-positive bacteria (*Staphylococcus aureus*), and fungi (*Pichia pastoris* and *Yarrowia lipolytica*) [33]. FREPs from *B. glabrata* are also able to bind trematode *Echinostoma paraensei* sporocysts, as well as a variety of microbes (Gram-positive and Gram-negative bacteria and yeast) with certain specificity with respect to pathogen type [29]. FREP4 of 65–75-kDa could bind to sporocysts and secretory/excretory products of *E. paraensei*, whereas FREPs with higher molecular-weight bound to Gram-positive, Gram-negative bacteria, and yeast [29]. Except for the binding activity, FREPs are also found to promote cellular immunity. The purified FREPs from serum of mussel *M. galloprovincialis* were reported to promote the phagocytosis of fluorescent beads [31].

However, the mechanisms of binding specificity of molluscan FREP and the polymorphism of Ig domain still need further investigation. The identification of more FREPs and their polymorphisms, and the comparison of structure and function will further clarify the taxonomic diversification of FREPs and their underlying recombination mechanisms in molluscs [20], providing clues to understanding the relationship between the diversification of FREP and specific immunity.

#### 4.1.2. Siglec and JAM-A Protein

Except for FREPs, several Ig domain only proteins have also been identified as vital PRRs in the mollusc immune system. Siglec is a member of IgSF, which involves in the host-pathogen recognition, cell–cell interactions and subsequent signaling pathways in the immune and nervous systems in vertebrates. The orthologs of mammalian Siglec are composed of one N-terminal V-set Ig domain, C2-set Ig domains, and cytosolic immunoreceptor tyrosine-based inhibitory motifs (ITIMs) [34,35]. The key arginine (Arg) residues in N terminal V-set Ig domain endow Siglecs with an ideal platform for sialic acid binding through a salt bridge [35]. Recently, a homolog of Siglec was identified in oyster *C. gigas* (CgSiglec-1) with two I-set Ig domains [36]. The recombinant CgSiglec-1 protein (rCgSiglec-1) could bind poly sialic acid (pSIAS), LPS, and PGN in a dose-dependent manner, which was different from other Siglecs characterized by binding ligands only in a sialic acid-dependent manner [36]. The significant difference might attribute to the N-terminal I-set Ig domain instead of the canonical V-set. The blockade of CgSiglec-1 by specific polyclonal antibodies enhanced the LPS-induced cell apoptosis and phagocytosis towards *V. splendidus*. The broader PAMP binding spectrum and the induction of cellular immunity might endow CgSiglec-1 with important roles in the immune response

of oyster against a variety of foreign invaders [36]. Junctional adhesion molecule (JAM), another example of IgSF with a couple of immunoglobulin domains, can act as a regulator in homeostasis and inflammation in vertebrates [37]. A structural homolog of JAM-A (designated CgJAM-A-L) was identified from oyster *C. gigas* with typical JAM-A domain, which functioned as an adhesion molecule [38]. The recombinant CgJAM-A-L could bind to LPS, PGN, LTA, MAN, GLU, and poly I:C, as well as various microorganisms including *M. luteus*, *S. aureus*, *E. coli*, *V. anguillarum*, *V. splendidus*, *P. pastoris*, and *Y. lipolytica*. The phagocytic rates of oyster hemocytes towards Gram-negative bacteria *V. anguillarum* and yeast *P. pastoris* were significantly enhanced after the incubation of rCgJAM-A-L [38]. These results collectively indicated that the Ig containing proteins in mollusc could function as recognition molecules and opsonins in the immune defense against invading pathogen. As the most primitive specie with homolog of Ig, the Ig containing proteins (such as CgSiglec and CgJAM-A) in mollusc would provide useful clues for the evolutionary study of Ig domains as well as the adaptive immunity.

#### 4.2. LRR Domain Containing Pattern Recognition Receptors

LRR domain is a most widespread protein domain, which has a conserved core consensus of L-x-x-L-x-L-x-x-N to form a  $\beta$ -strand followed by a more variable sequence [39]. PDB structures for LRR-containing proteins show the LRR domains in an arc or horseshoe shape [39]. The function of many LRR domains is to provide a structural framework for protein–protein interactions (PPIs) [39]. More than 60,000 LRR-containing proteins have been identified in viruses, bacteria, archaea, fungi, plants, and animals [40]. They are especially enriched in plants [16], invertebrates (e.g., sea urchin) [41], and Cephalochordata (e.g., amphioxus) [42], which lack adaptive immunity. In plants, only two kinds of PRRs have been reported, and all of them are composed of LRR domains [16]. In animals, LRR-containing proteins, such as Toll-like receptors (TLRs) and NOD-like receptors (NLRs), could recognize the molecular determinants from a structurally diverse collection of bacterial, fungal, viral and parasite-derived components via the LRR domain [43]. It is worth noting that the selectively assembling of LRR segments in variable lymphocyte receptor gene of jawless vertebrates generates structural diversity for antigen recognition, which builds a new alternative adaptive immune system [44].

The LRR-containing proteins could be divided into two groups, proteins containing LRR motif and other motifs or domains, and proteins merely containing LRR motif termed LRR-only proteins [45]. The TLRs are typical representatives for the former group, consisting of a solenoid-like LRRNT (LRR N-terminal)-LRR-LRRCT (LRR C-terminal) ectodomain for ligand recognition and a cytoplasmic TIR domain for signal transduction [46]. As a family of conserved PRRs, TLRs are widely distributed in nearly all animal phyla, and they are well studied in mollusc. So far, a series number of TLRs have been identified in molluscs and an expanded set of 83 TLR genes are annotated in the oyster genome, which can be divided into five groups based on the patterns of sequence relatedness [17]. Among them, sP-type (short protostome-like without LRRCT-LRRNT ectodomains) TLRs are the most greatly expanded PRRs in oyster [39], which are similar with equally extensive expanded P-type TLRs (protostome-like with LRRCT-LRRNT ectodomains) in *Drosophila* [47] and V-type (vertebrate-type) TLRs in the sea urchin [41], suggesting that species-specific TLR gene expansion is a frequent, independent occurrence in metazoan phylogeny. Moreover, the expanded oyster TLRs displayed a different expression profile induced by single or different combinations of bacteria/virus stimulation [17]. For example, the expression of TLR Cg26493-D2 elevated after challenges with LPS and all *Vibrio* strains, while Cg13671 was only induced by *V. anguillarum*. About ten bivalve TLRs have been identified to participate in immune recognition of various microbes and PAMPs. A newly identified non-phagocytic receptor CgTLR-6 from *C. gigas* exhibited affinity to LPS and PGN, and broader recognition spectrum for bacteria and fungi [48]. Astonishingly, CgTLR1–4 from *C. gigas* lack the activities to recognize specific PAMPs, but they can activate NF- $\kappa$ B responsive reporter in HEK293 cells with a dose-dependent manner [49], which supports the functional differentiation of

these primordial TLRs in molluscs. Moreover, three TLRs (*HcToll1*, *HcToll2*, and *HcToll3*) from the pearl mussel *H. cumingii* could respond to bacterial and viral challenge and regulate the expression of antimicrobial peptides (AMPs) [50–52]. A rather canonical MyD88-dependent TLR pathway was also demonstrated in scallop and this pathway was involved in immune signaling to activate downstream anti-oxidant, anti-bacteria reaction, and apoptosis [53]. Recently, 29 TLR signaling pathway components were identified in Zhikong scallop *Chlamys farreri*, including SARM (ORF08051) and TRAM/TICAM2 (ORF08726), indicating the possibility of MyD88-independent TLR signaling pathway in mollusc [54]. All of the evidences suggested that the ancient molluscan TLRs could recognize broader pathogens and related PAMPs, and mediate the downstream anti-bacteria reaction in a MyD88-dependent and MyD88-independent dependent pathway.

The versatile functions of proteins merely containing LRR motif (LRR-only proteins, LRRop) in immune response have also been recently reported in scallop *C. farreri*. Six CfLRRops (CfLRRop 1-6) containing varied number (1–4) of LRR domains have been identified [40,55,56] with different mRNA expression profiles and PAMP binding activities. The mRNA expression of CfLRRop-1, -2, and -3 in hemocytes could be induced by the stimulations of LPS, PGN, GLU, and poly I:C, while the mRNA transcripts of CfLRRop-4, -5, and -6 in hemocytes could respond to the stimulations of different microbes, including *V. anguillarum*, *M. luteus*, and *P. pastoris*. The recombinant CfLRRop-1, -3, -4, -5, -6 proteins, but not rCfLRRop-2 protein, could directly bind LPS, PGN, GLU, and poly I:C with varied affinity, indicating LRR-only proteins represented a new type of PRR in bivalve with similar recognition characters as TLR [40,55,56]. Additionally, rCfLRRop-1 and -2, but not rCfLRRop-3, could induce the release of TNF- $\alpha$  in primary cultured scallop hemocytes. rCfLRRop-4, -5, and -6 could induce varied immune effectors, including tumor necrosis factor  $\alpha$ , superoxide dismutase, catalase, and lysozyme. The versatile CfLRRops could function in immune response not only as PRRs, but also as an immune effector or pro-inflammatory factor, suggesting a functional differentiation among the diverse CfLRRops [40,55,56]. These preliminary results, together with future characterization of the large variety of unexplored LRRops in molluscs, will offer valuable information for understanding the comprehensive mechanism of immune recognition [40,55,56].

#### 4.3. Lectin Domain Containing Pattern Recognition Receptors

Lectins are the most classical carbohydrate-binding proteins to bind the glycans of glycoproteins and glycolipids with high affinity. They play crucial roles in the innate immune responses of invertebrates by recognizing and eliminating pathogens [57]. Based on the structure, animal lectins have been classified into at least 13 lectin families (Available online: [www.imperial.ac.uk/research/animallelectins/](http://www.imperial.ac.uk/research/animallelectins/)), and at least seven groups of lectins have been identified in mollusc, including C-type, P-type, F-type, I-type, galectins, ficolins, and chitinase-like lectins [58,59]. Two extracellular lectins including C-type lectins (CTLs) and galectins are discussed in this circumstance with the emphases on their conserved domains and functions, such as cell adhesion, cell signaling, glycoprotein clearance, and pathogen recognition in molluscs.

##### 4.3.1. C-Type Lectins (CTLs)

CTLs are structurally characterized by double-loop composed of two highly conserved disulfide bridges located at the bases of the loops, and they can recognize and bind to terminal sugars on glycoproteins and glycolipids in a calcium-dependent manner [58,60]. Generally, CTLs contain at least one carbohydrate-recognition domain (CRD) with about 120 residues. Presently, a more general term “C-lectin domain containing proteins” (CTLDCs) has been introduced to define the proteins with one or more CTLDs, regardless of carbohydrate or calcium binding ability [58]. CTLs are abundant in molluscs, and an expanded set of 266 CTLDCs were annotated in the genome of oyster *C. gigas* [17]. More than 30 CTLs have been identified in molluscs, one from *Venerupis philippinarum* (*VpCTL*) [61], two from abalone *Haliotis discus hannai* (*CLHd* and *HdhCTL1*) [62,63], *Hyriopsis cumingii* (*Hcperlucin* and *HcLec4*) [64,65], and Manila clams *Ruditapes philippinarum* (*MCL-3*

and MCL-4) [66–68], respectively, six from *C. farreri* (*CfLec-1-CfLec-4*, *CfLec-4b*, *CfLec-5*) [7], seven from oyster *C. gigas* (*CgCLec-1-CgCLec-7*) [69], and ten from bay scallop *A. irradians* (*AiCTL-1-AiCTL-10*) [7]. Accumulating evidences have favored that, CTLs in molluscs exhibit high diversity of structure and functions.

The diversity of CRD structure is mainly embodied in the number of CRD as well as variety of  $\text{Ca}^{2+}$ -binding site 2 in CRDs. Most molluscan CTLs contain single CRD, while there are three or four CRDs in *CfLec-3* and *CfLec-4* from *C. farreri* [70,71] and *AiCTL-9* from *A. irradians* [72], respectively. The architecture and phylogenetic analysis revealed that multi-domain CRDs in different lineages did not arise from a common multi-domain progenitor, and these proteins served distinct functions in different animal lineages [70]. The CRD contains a double-loop structure, and the second loop (also called long loop region) is involved in  $\text{Ca}^{2+}$ -dependent carbohydrate binding. There are four  $\text{Ca}^{2+}$ -binding sites in this structure, among which the site 2 is known to be involved in the carbohydrate binding [58,73,74]. For the  $\text{Ca}^{2+}$ -binding sites 2, there are two conserved motifs determining the CRD binding ability. In vertebrates, the first motif is always EPN (Glu-Pro-Asn) or QPD (Gln-Pro-Asp), and the second one always WND (Trp-Asn-Asp). While in invertebrates, much more diversities are reported, and there are at least seven types in the first motif and more than ten types in the second motif presented in different molluscan CTLs [59]. The first motif is considered as the key switch in the specificity of binding with carbohydrate, and the second one increases the binding affinity and specificity [73,75]. However, the information about the function of the second motif in molluscs is still limited, which is urgently needed for us to understand the mechanism of CTLs functioning as a PRR.

Diverse immunological functions of C-type lectins in molluscs have been demonstrated in the past decades. The mRNA transcripts of almost all the CTLs are increased after stimulation with pathogens or PAMPs, implying that they are involved in innate immune response [67,76]. As an important PRR, molluscan CTLs display high affinity to various PAMPs on the surface of pathogens, such as LPS from Gram-negative bacteria, PGN from Gram-positive bacteria, glucan and mannan from fungi, and so on. Especially, the clustering of multiple CRDs in one molecule showed different PAMPs binding spectrum and affinity, and thus endowed multi-CRD CTLs with a broader spectrum and higher affinity to bind PAMPs. For instance, *CfLec-3* from *C. farreri* with three CRDs can bind more PAMPs than *CfLec-1* and *CfLec-2* with single CRD [77]. *AiCTL-9*, a multi-CRD lectin from scallop *A. irradians*, can bind LPS, PGN, glucan, and mannan [72]. Among the recombinant proteins of four CRDs in *CfLec-4* from *C. farreri* (designated as rCRD1, rCRD2, rCRD3, and rCRD4), rCRD3 and rCRD4 could bind to all four PAMPs tested, while rCRD1 and rCRD2 could only bind to LPS and mannan [78]. It is noteworthy that molluscan CTLs with the same first motif of  $\text{Ca}^{2+}$ -binding site 2 displayed different PAMPs binding spectrums. In *C. farreri*, *CfLec-1* containing the motif EPD could bind LPS, PGN and mannan in vitro, while *CfLec-2* with the same motif could also bind zymosan besides the other three PAMPs [79,80]. Similarly, CRD1 of *HcLec4* in *H. cumingii* with a “LND” motif could bind to LPS and PGN, while CRD3 of *HcLec4* with “LND and EPN” motifs could not bind to LPS and PGN [64]. Besides nonself recognition, molluscan CTLs participate in innate immune responses, including agglutination, hemocyte phagocytosis, and encapsulation, and they even display bactericidal effect. For instance, most of molluscan CTLs exhibit agglutinating activity towards various bacteria and fungi [7]. The purified MCL, MCL-3, and MCL-4 from the plasma of Manila clam *R. philippinarum* not only markedly suppressed the growth of *Alteromonas haloplanktis*, but also significantly enhanced the hemocyte phagocytic ability toward the bacteria [67,81]. Similarly, the recombinant CTLs from *C. farreri* [79,80] and *C. gigas* [82] also inhibited the growth of several bacteria. In *H. cumingii*, the bacteria clearance efficiency and the expression level of AMPs were both increased after *HcLec4* was knocked down, revealing that *HcLec4* exerted its antibacterial effect by regulating the expression of AMPs at the early stage of bacterial infection [64].

A large number of CTLs identified from molluscs shared similar structural features and displayed versatile immunocompetence. Since molluscs lack of antibody mediated adaptive immunity,



abundant CTLs with diverse expression profiles and bioactivities might function as nonclonal effectors in the molluscan immune system.

#### 4.3.2. Galectin

Galectins are a family of most conserved and ubiquitous lectin, which is defined by a conserved CRD with a canonical amino acid sequence and affinity primary for  $\beta$ -galactoside [2]. Similar with C type lectins, some galectins have one single CRD, while other galectins contain a clustering of multiple CRDs, such as *CvGal* identified from the hemocytes of *C. virginica* [83], *PoGal-1* and *PoGal-2* from *P. fucata* [84,85], *AiGal-1* and *AiGal-2* from *A. irradians* [86,87], and *MCGal* from clam *R. philippinarum* [88]. The clustering of multiple CRDs organization was uniquely found in molluscan galectins, in which different CRDs might share a common ancestor, according to the phylogenetic analysis [83,85].

It has been demonstrated that galectin of molluscs can target glycans on the surfaces of bacteria and parasites, and plays a crucial role in innate immune responses [7]. Two galectin in *P. fucata*, *PoGal-1* and *PoGal-2*, were significantly up-regulated after *V. alginolyticus* stimulation, suggesting that *PoGals* were involved in the immune response against bacteria [84,85]. *CvGal* in *C. virginica* facilitated particularly the recognition of a variety of potential microbial pathogens, unicellular algae, and preferentially *Perkinsus* trophozoites [83]. The strong interaction between *CvGal* and *P. trophozoites* suggests that *CvGal* functions as a hemocyte surface receptor for the parasite and mediates phagocytosis [83]. *MCGal* exhibited an affinity towards galactose and N acetylgalactosamine and could bind to the surface of *Perkinsus olseni* and agglutinate *V. tapetis* in vitro [88]. *AiGal-2* from *A. irradians* displayed high binding and agglutinating activities toward a variety of microbes, including *V. anguillarum*, *M. luteus*, *E. coli*, *V. fluvialis*, and *Edwardsiella tarda*, and could also recruit hemocytes and promote hemocytes encapsulation. These results suggested that *AiGal2* functioned as a pattern recognition receptor in immune defense and contributed to the nonself recognition and elimination in cellular immune response of bay scallop [87].

#### 4.3.3. Other Lectins

F-type lectins (fucolectins) are fucose-binding proteins, which have also been identified in oyster *Pinctada martensii* (*PmF*-lectin) and *C. gigas*. The mRNA transcript of *PmF*-lectin is abundant in hemocytes and gill, and its expression level is dramatically up-regulated after a challenge with *V. alginolyticus*, suggesting that *PmF*-lectin is involved in the innate immune response [89].

The chitinase-like lectins and ficolin-like proteins are also found in molluscs. Two chitinase-like proteins (*CgClp1* and *CgClp2*) were identified from *C. gigas*, and their mRNA expressions in hemocytes could be induced by LPS stimulation, strongly suggesting that these two close paralogous genes were involved in oyster immunity [90]. A novel ficolin-like protein (*ChFCN*) was identified from *C. hongkongensis*. It contained a typical signal peptide and a fibrinogen-related domain at the N- and C-terminus, respectively, but no additional ficolin collagen-like domain. *ChFCN* mRNA transcripts increased under microbial infections and the recombinant *ChFCN* protein could bind and agglutinate microorganisms, and enhance the phagocytosis [91].

Recently, some new family lectins were proposed in molluscs. For instance, a glycosylated galactose-binding lectin (*MCL*) from *M. californianus*, showed antibacterial activities against Gram-positive and Gram-negative bacteria, but didn't belong to any previously reported lectin family via bioinformatics analysis [92]. A new lectin from *Aplysia dactylomela* eggs (*ADEL*) was isolated by affinity chromatography with the activities to agglutinate and inhibit biofilm formation of *Staphylococcus aureus*, suggesting that this lectin might be a potential alternative to the conventional use of antimicrobial agents in the treatment of infections caused by *Staphylococcal* biofilms [93].

#### 4.4. C1q Domain Containing Pattern Recognition Receptors

C1q domain containing proteins (C1qDCs) are a family of proteins containing a globular head C1q domain (ghC1q) in C-terminus, which recognize a broad range of ligands and trigger a serial of immune responses, such as the classical pathway of complement [94]. There were several C1qDC proteins identified from molluscs, including *M. galloprovincialis* (MgC1q) [95], pearl mussel *Hyriopsis cumingii* (HcC1qDC5) [96], *C. farreri* (CfC1qDC and CfC1qDC-2), *A. irradians* (AiC1qDC-1 and AiC1qDC-2) [7], and *C. gigas* (CgC1qDC-1–4) [7]. All of the molluscan C1qDCs contain a typical gC1q domain with a 10-stranded  $\beta$ -sandwich fold forming a jelly-roll topology, and they are found to exert diverse biological functions in molluscan immunity. For instance, the mRNA expression level of MgC1q in *M. galloprovincialis* hemocytes was drastically increased after the stimulations of Gram-positive bacteria and Gram-negative bacteria [95]. The recombinant CfC1qDC (rCfC1qDC) and AiC1qDC-2 (rAiC1qDC-2) proteins both bound LPS, PGN,  $\beta$ -glucan, and poly I:C, while the latter could also bind mannan and yeast-glucan [97,98]. The broader PAMP binding activities of AiC1qDC-2 contributed to its broader bacterial agglutinating spectrum to fungus *P. pastoris*, Gram-positive bacteria *B. subtilis*, and Gram-negative bacteria *E. coli* and *V. anguillarum* [99]. Moreover, CgC1qDCs in oyster (CgC1qDC-2–4) could enhance the phagocytosis of oyster hemocytes, and the enhancements towards Gram-negative bacteria were significantly higher than that towards Gram-positive bacteria, indicating the important roles of molluscan C1qDC as an opsonin in the clearance of invading microbes [100]. Comparatively, CgC1qDC-3 exhibited higher binding affinity to LPS and stronger opsonization, and its mRNA expression could be up-regulated rapidly and persisted upon the secondary challenge with homologous *Vibrios*. It was proposed to exert efficient functions in the immune response against invading pathogens [100]. Most interestingly, CfC1qDC could interact with human heat-aggregated IgG, which gave a hint for the conservative property of C1qDC proteins to functionally interact with immunoglobulins in evolution [98].

An expanded set of 321 and 168 C1qDCs are annotated in oyster [17] and mussel genome [95], which are promising molecules that are involved in immune specificity with functional divergence and provide clues for the knowledge of evolution of the complement system.

#### 4.5. SR Domain Containing Pattern Recognition Receptors

SRs represent a large family of endocytic receptors with multifunction to bind and internalize a variety of microbial pathogens, as well as modified or endogenous molecules derived from the host, and contribute to a range of physiological or pathological processes [101]. According to their multi-domain structure, SRs are classified into six different classes (classes A, B, C, D, E, and F) [102]. However, the information about molluscan SRs is extremely limited, and only two SRs have been identified in oyster *P. martensii* and scallop *C. farreri* [103]. The mRNA expression of CfSR was significantly up-regulated by the stimulations of LPS, PGN, and  $\beta$ -glucan. The recombined protein CfSR displayed unique broad ligand-binding properties not only for acetylated low density lipoprotein (Ac-LDL) and dextran sulfate, but also for various PAMPs, such as LPS, PGN, mannan, and zymosan [103]. Moreover, the increasing genomic information indicates that there is an expansion of SRs and about 71 members have been annotated in oyster, and more especially, they have experienced high levels of diversification [17]. The mRNA transcripts of SRs in oyster *C. gigas* are highly up-regulated in response to summer mortality syndrome, hypoxia, and bacterial challenge [69], suggesting their significant roles in adaption to complex environment.

#### 4.6. Other Domain Containing Pattern Recognition Receptors

##### 4.6.1. Peptidoglycans Recognition Proteins (PGRPs)

PGRPs are a kind of highly conserved PRRs that specifically bind to PGN, a unique cell wall component of all virtual bacteria, but is not present in eukaryotic cells [104]. Nine PGRP genes are predicated in the oyster *C. gigas* genome, and six of them (CgPGRP-S1S, -S1L, -S2, -S3, -S4,

and CgPGRP-L) have been characterized [17,69]. A positive Darwinian selection was suggested by the phylogenetic analysis in the CgPGRP family [105]. Molluscan PGRPs have also been identified from the bay scallop *A. irradians* (AiPGRP) [106] and Zhikong scallop *C. farreri* (CfPGRP S1) [107]. The identified molluscan PGRPs are all short type with a conserved amidase-2/PGRP domain in their C-terminus, suggesting that these proteins retain the conserved amidase activity. Although the molecular weight of CgPGRP-L is close to that of long PGRP groups, it is homologous to short PGRPs and possesses an additional goose-type (g-type) lysozyme domain, suggesting that CgPGRP-L may have both binding and lytic functions against the bacterial cell wall [108]. CfPGRP-S1 displayed strong activities to agglutinate Gram-positive bacteria *M. luteus* and *Bacillus subtilis*, while weak activities to agglutinate Gram-negative bacteria *E. coli* [107]. It also functioned as a zinc-dependent amidase to degrade PGN and strongly inhibited the growth of *E. coli* and *S. aureus* in virtue of the conserved Zn<sup>2+</sup> binding sites in amidase-2/PGRP domain. The recombinant PGRP-S1S (rPGRP-S1S) from oyster *C. gigas* did not show any bactericidal activity, but agglutinated *E. coli* and induced secretion of granular contents by hemocyte degranulation [109]. Although all of the molluscan PGRP genes contain conserved PGRP/amidase domain, the critical residues involved in specific PGN recognition show a certain degree of mutation, which may create a more flexible response to different microbial challenges.

#### 4.6.2. Gram-Negative Bacteria-Binding Proteins (GNBPs)

The GGBP family, including the LPS and  $\beta$ -1,3-glucan binding protein (LGBP), and  $\beta$ -1,3-Glucan binding proteins ( $\beta$ GBP), can bind Gram-negative bacteria, LPS, and GLU. The information about molluscan GGBP is still limited and there are only a few GGBP genes identified from *P. fucata* (poLGBP) [110] and *C. farreri* (CfLGBP) [111], respectively. They both possess a potential polysaccharide-binding motif, a glucanase motif, an LPS-binding site, and a  $\beta$ -1,3-linkage of polysaccharide. Their mRNA expressions were significantly up-regulated after bacteria or LPS stimulation, indicating a possibly important role during bacterial infection [110,111]. The CfLGBP exhibited obvious binding and agglutination activity toward various PAMPs and microorganisms, and the polymorphism of 7679 G/G in CfLGBP enhanced its binding activity to LPS and  $\beta$ -glucan, which was associated with the disease resistance of scallop against *V. anguillarum* [112]. There are two cDNA fragments coding  $\beta$ GBPs (Cg $\beta$ GBP-1 and Cg $\beta$ GBP-2) reported in oyster. Although domain structures of both Cg $\beta$ GBPs are similar to other invertebrate  $\beta$ GBPs, they seem to have evolved with different immunological functions. The recombinant Cg $\beta$ GBP-2 enhanced the phenoloxidase (PO) activity of hemocyte suspensions under the presence of laminarin, but rCg $\beta$ GBP-1 did not show this enhancement. Because there are integrin recognition sites that are identified in Cg $\beta$ GBP-1, but not in Cg $\beta$ GBP-2, it is suggested that the hemocyte-related functions of Cg $\beta$ GBP-1 are possibly evolved from integrin [112]. Additionally, a  $\beta$ GBP was purified from the plasma of marine mussel *Perna viridis* with an inherent serine protease activity [113]. It agglutinated baker yeast, bacteria, and erythrocytes, and enhanced pro-PO activity of the plasma. Although the gene sequence and molecular structure are still unknown, molluscan  $\beta$ GBPs are considered to function as recognition molecule for  $\beta$ -1,3-glucan, and also mediation molecule in the immune response such as PO activation in molluscs.

### 5. Novel Carbohydrates Recognition Protein in Molluscs

Except for the canonical carbohydrates recognition proteins, some novel molecules were discovered in molluscs to play roles in PAMPs recognition. Concomitantly, some unconventional immune defense strategies or novel pathopoiesis mechanisms are prompted with great potentiality. These un-traditional recognition domains or proteins are under revealing with increasing attention.

#### 5.1. Novel DM9 Domain Typically Recognizing Mannose

Recently, a novel PRR with two DM9 domains was identified from oyster *C. gigas* (CgDM9CP-1) by two laboratories. DM9 domain was originally identified in *Drosophila melanogaster* with no

defined functions. The DM9 containing protein (CgDM9CP-1) was significantly enriched from oyster hemolymph by affinity chromatography through a mannose-conjugated cellulose column. CgDM9CP-1 exhibited high binding specificity and avidity toward d-mannose residue [114]. It served as a PRR with a broad range of recognition spectrum to various PAMPs, including LPS, PGN, MAN, and  $\beta$ -1,3-glucan in a d-mannose-dependent manner, as well as bacteria and fungi. The crystal structures of wild-type and loss-of-function mutant of CgDM9CP-1 demonstrated that Asp22 and Lys43 were the critical residues for ligand recognition [114,115]. Moreover, CgDM9CP-1 protein was found to mainly distribute on the surface of *C. gigas* hemocytes, and it could be translocated into cytoplasm and colocalized with the engulfed microbes during hemocytes phagocytosis, which collectively indicated its important role as classical PRR. It is worth noting that DM9 domain has been found to exist in various proteins from a number of species, and its existence in prokaryotic cells indicates that the universal DM9 domain is an ancient protein domain that probably evolved from prokaryotes [114]. But, the conserved functions of DM9CPs domain during evolution and their intracellular signal pathways still need further investigation. A total of 477 DM9CPs have been annotated in organisms from the Prokaryotae, Fungi, Protista, and Animalia Kingdoms, but just not Plantae Kingdom. The uneven distribution of phylogenetic patterns of the DM9 domain is likely to reflect the natural selection during molecular evolution in innate immunity.

### 5.2. Caspase with Specific Binding Activity to LPS

Caspases is well-known as a family of cysteine-dependent aspartate-directed proteases for their key roles in apoptosis and inflammatory responses. Recently, they were reported as the PRRs to play important role in immune recognition. Following an unexpected result that cytoplasmic LPS could activate caspase-11 “non-canonical inflammasome” in a TLR4-independent manner in mice [116], human caspase-4, -5, and mouse homologue caspase-11 were found to directly bind intracellular LPS by their CARD domain with high specificity and affinity. The intracellular LPS induced the oligomerization and activation of caspase-4/5/11, leading to pyroptosis [117]. The recognition character of caspase to LPS was favored in lower invertebrate oyster *C. gigas* sooner after the first publication. The oyster CgCaspase-3 was confirmed to have the specific binding activity to LPS, but not to LTA, mannan, and  $\beta$ -1,3-glucan by ELISA-based LPS binding assay and surface plasmon resonance (SPR) analysis. The *in vivo* interaction of CgCaspase-3 with LPS specifically inhibited the cell apoptosis, just opposite to pyroptosis promoting effect of LPS-caspase binding in mammals [118]. Interestingly, the LPS binding mechanism of CgCaspase-3 was dependent on CASc domain, which was totally different from mammalian caspase-4/5/11 relying on the CARD domain. In addition, CgCaspase-1, only containing one CASc domain, also displayed direct LPS binding activity, which was similar to CgCaspase-3 [119]. The great evolutionary difference between mammalian and molluscan caspases encourages us to suspect that molluscs might evolve sophisticated strategies in both LPS binding mechanism and the following immune effect. It seems that the CARD-dependent LPS binding activity of mammalian caspase mainly promotes pyroptosis, while CASc-dependent LPS binding activity of molluscs caspase inhibits apoptosis. But, the detailed mechanisms and the functional differentiation among all the caspases harboring CARD or CAS domain still need further investigation.

### 5.3. IL-17 with Lectin and Bactericidal Activity

In mammals, cytokines are the major regulators of the host immune response by the interaction with specific receptors. An unexpected carbohydrate-binding (lectin) property of cytokines was discovered in the study of cytokine-receptor interaction. Some cytokines could recognize and bind to the specific glycosylation sites of their individual receptors and induce immune modulation [120,121]. For example, recombinant human interleukins IL-1 $\alpha$ , IL-1 $\beta$ , IL-4, IL-6, and IL-7 exhibited different and specific calcium-independent carbohydrate-binding properties and such activities were essential for providing specific signaling systems [122,123]. The carbohydrate-recognition domains (CRDs) localized at the opposite of the receptor-binding domain endow these interleukin molecules with

bi-functional regulation property [120]. However, opposite to this kind of endogenous carbohydrate binding activity, one ancient invertebrate interleukin displayed heterogenous PAMPs binding activity. The recombinant IL17-5 from oyster *C. gigas* (rCgIL17-5) was proved to directly bind to PGN, LPS, poly (I:C), and  $\beta$ -1,3-glucan, with the highest affinity to PGN, which had never been reported in vertebrate interleukins [124]. The mRNA expression level of rCgIL17-5 was strongly up-regulated after PGN stimulation, suggesting that its involvement in the recognition of diverse pathogens as a PRR and preference to recognize Gram-positive bacteria [124]. Another lectin like cytokine in earthworm *Eisenia foetida*, the coelomic cytolytic factor (CCF), displayed both lectin activity to bind LPS, PGN, and GLU, as well as TNF-like lytic activity [125]. The lectin-like activity/domain of IL and CCF may represent an essential recognition mechanism that has been functionally conserved during the innate immune response of invertebrates and vertebrates [126]. Moreover, rCgIL17-5 could significantly inhibit the growth of both Gram-positive bacteria *M. luteus* and Gram-negative bacteria *E. coli*, indicating its direct anti-bacteria effect by inhibiting their proliferation [124]. The endogenous carbohydrate binding activity of vertebrate cytokine and the heterogenous PAMPs binding activity of invertebrate cytokine were suspected to be a result of functional differentiation during evolution [124].

#### 5.4. Metabolic Enzymes with New Role of Carbohydrate Binding

Metabolic enzymes have long been known accustomedly for their function of catalyzing substrate to produce new products. Recently, the unexpected non-metabolic functions of metabolic enzymes were observed in the context of inflammation in both vertebrates and invertebrates. For example, the metabolic enzyme hexokinase could unexpectedly act as a pattern recognition receptor to recognize bacterial peptidoglycan and trigger the activation of inflammasome in mice [127,128]. Similarly, a phosphoenolpyruvate carboxykinase (CgPEPCK) in oyster *C. gigas*, which is well known as a key enzyme that is involved in the metabolic pathway of gluconeogenesis, also displays the LPS and PGN binding activities [129]. Besides, arginine kinase, a conserved essential phosphagen kinase (PK) in ATP buffering systems, has also been identified (designated CgAK) from the hemolymph of *C. gigas* by LPS affinity chromatography. CgAK could directly bind to LPS in a concentration-dependent manner. The interaction with LPS significantly decreased the ATP hydrolytic activity of CgAK and this in turn led to the accumulation of ATP in vitro, which is of great importance for the detection of pathogens and the induction of host immune responses [130]. These new progresses have subverted our traditional awareness of metabolic enzymes and inferred that their overlooked direct immune functions are evolutionarily conserved in both vertebrates and invertebrates. The latest progress in invertebrates could be also referenced for new therapeutic strategies development by exploiting these non-canonical features of metabolic machinery, modulating their contribution to the immune response without impacting their basal metabolic functions [128].

## 6. Conclusions

Recognition is the fundamental step for the interaction of hosts–pathogens. Many glycans show remarkably discontinuous distribution across various microbes, and certain lineage-specific glycans have become important signals for nonself recognition of hosts by a toolkit of innate recognizing molecules [131]. Coincidentally, the recognition domains in multicellular organisms, especially in invertebrates, also show significantly species-specific expansion in genome [17]. A lot of members of the expanded PRR families have been confirmed to have versatile functions in immune response, and their functional cooperation and differentiation might provide a subtle discrimination of self from nonself. It could be inferred that there might be a correlation between the lineage-specific glycan of pathogens and species-specific expansion of innate recognition domains in host. However, the detailed mechanism about the species-specific interaction between versatile innate recognition receptors with some common domains and microbes need much more researches. In molluscs, the great biodiversity of species provide plenty of solutions for pathogen recognition, and the presence of open circulatory system and ancient specialized hemocytes offer

effective immunological surveillance. A systemic comparison of the heterogeneous PRRs might be needed for better understanding their functional differentiation. Meanwhile, with the rapid development of multi-omics biotechnology, the comprehensive and systematic understanding about the carbohydrate recognition during host-microbes interaction and the other processes might be revealed by big data analysis of genome information in hosts and great glycomic data in pathogens.

**Acknowledgments:** This research was supported by a grant from National Science Foundation of China (No. 31530069), AoShan Talents Cultivation Program Supported by Qingdao National Laboratory for Marine Science and Technology (No. 2017ASTCP-OS13), earmarked funds from Modern Agro-industry Technology Research System (CARS-49), Dalian High Level Talent Innovation Support Program (2015R020), Talented Scholars in Dalian Ocean University, and Natural Science Foundation of Liaoning, China (20170520056).

**Author Contributions:** All authors contributed to preparation of the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Medzhitov, R.; Janeway, C., Jr. Innate immune recognition: Mechanisms and pathways. *Immunol. Rev.* **2000**, *173*, 89–97. [[CrossRef](#)] [[PubMed](#)]
2. Janeway, C.A., Jr.; Medzhitov, R. Innate immune recognition. *Annu. Rev. Immunol.* **2002**, *20*, 197–216. [[CrossRef](#)] [[PubMed](#)]
3. Pasmán, L.; Kasper, D.L. Building conventions for unconventional lymphocytes. *Immunol. Rev.* **2017**, *279*, 52–62. [[CrossRef](#)] [[PubMed](#)]
4. Spoel, S.H.; Dong, X. How do plants achieve immunity? Defence without specialized immune cells. *Nat. Rev. Immunol.* **2012**, *12*, 89–100. [[CrossRef](#)] [[PubMed](#)]
5. McMahon, B.R.; Wilkens, J.L.; Smith, P.J.S. Invertebrate Circulatory Systems. In *Handbook of Physiology, Comparative Physiology*; Dantzler, W.H., Ed.; Oxford University Press: New York, NY, USA, 2011; pp. 931–1008.
6. Dunn, C.W.; Giribet, G.; Edgecombe, G.D.; Hejnol, A. Animal Phylogeny and Its Evolutionary Implications. *Annu. Rev. Ecol. Evol. Syst.* **2014**, *45*, 371–395. [[CrossRef](#)]
7. Song, L.; Wang, L.; Zhang, H.; Wang, M. The immune system and its modulation mechanism in scallop. *Fish Shellfish Immun.* **2015**, *46*, 65–78. [[CrossRef](#)] [[PubMed](#)]
8. Buckeridge, M.S. The evolution of the Glycomic Codes of extracellular matrices. *BioSystems* **2018**, *164*, 112–120. [[CrossRef](#)] [[PubMed](#)]
9. Medzhitov, R. Recognition of microorganisms and activation of the immune response. *Nature* **2007**, *449*, 819–826. [[CrossRef](#)] [[PubMed](#)]
10. Gabius, H.J. The sugar code: Why glycans are so important. *BioSystems* **2018**, *164*, 102–111. [[CrossRef](#)] [[PubMed](#)]
11. Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27*, 3–49. [[CrossRef](#)] [[PubMed](#)]
12. Rudd, P.M.; Wormald, M.R.; Dwek, R.A. Sugar-mediated ligand-receptor interactions in the immune system. *Trends Biotechnol.* **2004**, *22*, 524–530. [[CrossRef](#)] [[PubMed](#)]
13. Morelli, L.; Poletti, L.; Lay, L. Carbohydrates and Immunology: Synthetic Oligosaccharide Antigens for Vaccine Formulation. *Eur. J. Org. Chem.* **2011**, *2011*, 5723–5777. [[CrossRef](#)]
14. Imberty, A.; Varrot, A. Microbial recognition of human cell surface glycoconjugates. *Curr. Opin. Struct. Biol.* **2008**, *18*, 567–576. [[CrossRef](#)] [[PubMed](#)]
15. Palssonmcdermott, E.M.; O'Neill, L.A.J. Building an immune system from nine domains. *Biochem. Soc. Trans.* **2007**, *35*, 1437–1444. [[CrossRef](#)] [[PubMed](#)]
16. Padmanabhan, M.; Cournoyer, P.; Dineshkumar, S.P. The leucine-rich repeat domain in plant innate immunity: A wealth of possibilities. *Cell. Microbiol.* **2009**, *11*, 191–198. [[CrossRef](#)] [[PubMed](#)]
17. Zhang, L.; Li, L.; Guo, X.; Litman, G.W.; Dishaw, L.J.; Zhang, G. Massive expansion and functional divergence of innate immune genes in a protostome. *Sci. Rep.* **2015**, *5*, 8693. [[CrossRef](#)] [[PubMed](#)]
18. Rodrigueznunez, I.; Wcisel, D.J.; Litman, G.W.; Yoder, J.A. Multigene families of immunoglobulin domain-containing innate immune receptors in zebrafish: Deciphering the differences. *Dev. Comp. Immunol.* **2014**, *46*, 24–34. [[CrossRef](#)] [[PubMed](#)]

19. Zhang, L.; Li, L.; Zhang, G. Sequence variability of fibrinogen-related proteins (FREPs) in *Crassostrea gigas*. *Chin. Sci. Bull.* **2012**, *57*, 3312–3319. [[CrossRef](#)]
20. Adema, C.M. Fibrinogen-Related Proteins (FREPs) in Mollusks. *Results Probl. Cell Differ.* **2015**, *57*, 111–129. [[PubMed](#)]
21. Cerenius, L.; Soderhall, K. Variable immune molecules in invertebrates. *J. Exp. Biol.* **2013**, *216*, 4313–4319. [[CrossRef](#)] [[PubMed](#)]
22. Leonard, P.M.; Adema, C.M.; Zhang, S.; Loker, E.S. Structure of two FREP genes that combine IgSF and fibrinogen domains, with comments on diversity of the FREP gene family in the snail *Biomphalaria glabrata*. *Gene* **2001**, *269*, 155–165. [[CrossRef](#)]
23. Huang, B.; Zhang, L.; Li, L.; Tang, X.; Zhang, G. Highly diverse fibrinogen-related proteins in the Pacific oyster *Crassostrea gigas*. *Fish Shellfish Immun.* **2015**, *43*, 485–490. [[CrossRef](#)] [[PubMed](#)]
24. Zhang, S.; Nian, H.; Zeng, Y.; Dejong, R.J. Fibrinogen-bearing protein genes in the snail *Biomphalaria glabrata*: Characterization of two novel genes and expression studies during ontogenesis and trematode infection. *Dev. Comp. Immunol.* **2008**, *32*, 1119–1130. [[CrossRef](#)] [[PubMed](#)]
25. Zhang, S.; Loker, E.S. Representation of an immune responsive gene family encoding fibrinogen-related proteins in the freshwater mollusc *Biomphalaria glabrata*, an intermediate host for *Schistosoma mansoni*. *Gene* **2004**, *341*, 255–266. [[CrossRef](#)] [[PubMed](#)]
26. Gorbushin, A.M.; Panchin, Y.V.; Iakovleva, N.V. In search of the origin of FREPs: Characterization of *Aplysia californica* fibrinogen-related proteins. *Dev. Comp. Immunol.* **2010**, *34*, 465–473. [[CrossRef](#)] [[PubMed](#)]
27. Simakov, O.; Marletaz, F.; Cho, S.; Edsingeronzales, E.; Havlak, P.; Hellsten, U.; Kuo, D.; Larsson, T.; Lv, J.; Arendt, D.; et al. Insights into bilaterian evolution from three spiralian genomes. *Nature* **2012**, *493*, 526–531. [[CrossRef](#)] [[PubMed](#)]
28. Zhang, S.; Adema, C.M.; Kepler, T.B.; Loker, E.S. Diversification of Ig Superfamily Genes in an Invertebrate. *Science* **2004**, *305*, 251–254. [[CrossRef](#)] [[PubMed](#)]
29. Zhang, S.; Zeng, Y.; Loker, E.S. Expression profiling and binding properties of fibrinogen-related proteins (FREPs), plasma proteins from the schistosome snail host *Biomphalaria glabrata*. *Innate Immun.* **2008**, *14*, 175–189. [[CrossRef](#)] [[PubMed](#)]
30. Hanington, P.C.; Zhang, S. The Primary Role of Fibrinogen-Related Proteins in Invertebrates Is Defense, Not Coagulation. *J. Innate Immun.* **2011**, *3*, 17–27. [[CrossRef](#)] [[PubMed](#)]
31. Romero, A.; Dios, S.; Poisabeiro, L.; Costa, M.M.; Posada, D.; Figueras, A.; Novoa, B. Individual sequence variability and functional activities of fibrinogen-related proteins (FREPs) in the Mediterranean mussel (*Mytilus galloprovincialis*) suggest ancient and complex immune recognition models in invertebrates. *Dev. Comp. Immunol.* **2011**, *35*, 334–344. [[CrossRef](#)] [[PubMed](#)]
32. Zhang, H.; Wang, L.; Song, L.; Song, X.; Wang, B.; Mu, C.; Zhang, Y. A fibrinogen-related protein from bay scallop *Argopecten irradians* involved in innate immunity as pattern recognition receptor. *Fish Shellfish Immun.* **2009**, *26*, 56–64. [[CrossRef](#)] [[PubMed](#)]
33. Yang, C.; Wang, L.; Zhang, H.; Wang, L.; Huang, M.; Sun, Z.; Sun, Y.; Song, L. A new fibrinogen-related protein from *Argopecten irradians* (AiFREP-2) with broad recognition spectrum and bacteria agglutination activity. *Fish Shellfish Immun.* **2014**, *38*, 221–229. [[CrossRef](#)] [[PubMed](#)]
34. Varki, A. Natural ligands for CD33-related Siglecs? *Glycobiology* **2009**, *19*, 810–812. [[CrossRef](#)] [[PubMed](#)]
35. Brinkman-Van der Linden, E.C.; Hurtado-Ziola, N.; Hayakawa, T.; Wiggleton, L.; Benirschke, K.; Varki, A.; Varki, N. Human-specific expression of Siglec-6 in the placenta. *Glycobiology* **2007**, *17*, 922–931. [[CrossRef](#)] [[PubMed](#)]
36. Liu, C.; Jiang, S.; Wang, M.; Wang, L.; Chen, H.; Xu, J.; Lv, Z.; Song, L. A novel siglec (CgSiglec-1) from the Pacific oyster (*Crassostrea gigas*) with broad recognition spectrum and inhibitory activity to apoptosis, phagocytosis and cytokine release. *Dev. Comp. Immunol.* **2016**, *61*, 136–144. [[CrossRef](#)] [[PubMed](#)]
37. Ludwig, R.J.; Hardt, K.; Hatting, M.; Bistrrian, R.; Diehl, S.; Radeke, H.H.; Podda, M.; Schon, M.P.; Kaufmann, R.; Henschler, R.; et al. Junctional adhesion molecule (JAM)-B supports lymphocyte rolling and adhesion through interaction with  $\alpha 4\beta 1$  integrin. *Immunology* **2009**, *128*, 196–205. [[CrossRef](#)] [[PubMed](#)]
38. Liu, C.; Wang, M.; Jiang, S.; Wang, L.; Chen, H.; Liu, Z.; Qiu, L.; Song, L. A novel junctional adhesion molecule A (CgJAM-A-L) from oyster (*Crassostrea gigas*) functions as pattern recognition receptor and opsonin. *Dev. Comp. Immunol.* **2016**, *55*, 211–220. [[CrossRef](#)] [[PubMed](#)]

39. Ng, A.; Eisenberg, J.; Heath, R.J.; Huett, A.; Robinson, C.M.; Nau, G.J.; Xavier, R.J. Human leucine-rich repeat proteins: A genome-wide bioinformatic categorization and functional analysis in innate immunity. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 4631–4638. [[CrossRef](#)] [[PubMed](#)]
40. Wang, M.; Wang, L.; Guo, Y.; Yi, Q.; Song, L. An LRR-only protein representing a new type of pattern recognition receptor in *Chlamys farreri*. *Dev. Comp. Immunol.* **2016**, *54*, 145–155. [[CrossRef](#)] [[PubMed](#)]
41. Rast, J.P.; Smith, L.C.; Lozacoll, M.; Hibino, T.; Litman, G.W. Genomic Insights into the Immune System of the Sea Urchin. *Science* **2006**, *314*, 952–956. [[CrossRef](#)] [[PubMed](#)]
42. Dishaw, L.J.; Haire, R.N.; Litman, G.W. The amphioxus genome provides unique insight into the evolution of immunity. *Brief Funct. Genom.* **2012**, *11*, 167–176. [[CrossRef](#)] [[PubMed](#)]
43. Ng, A.; Xavier, R.J. Leucine-rich repeat (LRR) proteins: Integrators of pattern recognition and signaling in immunity. *Autophagy* **2011**, *7*, 1082–1084. [[CrossRef](#)] [[PubMed](#)]
44. Herrin, B.R.; Cooper, M.D. Alternative Adaptive Immunity in Jawless Vertebrates. *J. Immunol.* **2010**, *185*, 1367–1374. [[CrossRef](#)] [[PubMed](#)]
45. Huang, S.; Yuan, S.; Guo, L.; Yu, Y.; Li, J.; Wu, T.; Liu, T.; Yang, M.; Wu, K.; Liu, H. Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. *Genome Res.* **2008**, *18*, 1112–1126. [[CrossRef](#)] [[PubMed](#)]
46. Takeda, K.; Akira, S. Toll-like receptors in innate immunity. *Int. Immunol.* **2004**, *17*, 1–14. [[CrossRef](#)] [[PubMed](#)]
47. Rock, F.L.; Hardiman, G.; Timans, J.C.; Kastelein, R.A.; Bazan, J.F. A family of human receptors structurally related to *Drosophila* Toll. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 588–593. [[CrossRef](#)] [[PubMed](#)]
48. Wang, W.; Zhang, T.; Wang, L.; Xu, J.; Li, M.; Zhang, A.; Qiu, L.; Song, L. A new non-phagocytic TLR6 with broad recognition ligands from Pacific oyster *Crassostrea gigas*. *Dev. Comp. Immunol.* **2016**, *65*, 182–190. [[CrossRef](#)] [[PubMed](#)]
49. Zhang, Y.; He, X.; Yu, F.; Xiang, Z.; Li, J.; Thorpe, K.L.; Yu, Z. Characteristic and Functional Analysis of Toll-like Receptors (TLRs) in the lophotrocozoan, *Crassostrea gigas*, Reveals Ancient Origin of TLR-Mediated Innate Immunity. *PLoS ONE* **2013**, *8*, e76464.
50. Ren, Q.; Zhong, X.; Yin, S.-W.; Hao, F.-Y.; Hui, K.-M.; Zhang, Z.; Zhang, C.-Y.; Yu, X.-Q.; Wang, W. The first Toll receptor from the triangle-shell pearl mussel *Hyriopsis cumingii*. *Fish Shellfish Immun.* **2013**, *34*, 1287–1293. [[CrossRef](#)] [[PubMed](#)]
51. Ren, Q.; Lan, J.-F.; Zhong, X.; Song, X.-J.; Ma, F.; Hui, K.-M.; Wang, W.; Yu, X.-Q.; Wang, J.-X. A novel Toll like receptor with two TIR domains (*HcToll-2*) is involved in regulation of antimicrobial peptide gene expression of *Hyriopsis cumingii*. *Dev. Comp. Immunol.* **2014**, *45*, 198–208. [[CrossRef](#)] [[PubMed](#)]
52. Zhang, H.-W.; Huang, Y.; Man, X.; Wang, Y.; Hui, K.-M.; Yin, S.-W.; Zhang, X.-W. *HcToll3* was involved in anti-*Vibrio* defense in freshwater pearl mussel, *Hyriopsis cumingii*. *Fish Shellfish Immun.* **2017**, *63*, 189–195. [[CrossRef](#)] [[PubMed](#)]
53. Wang, M.; Yang, J.; Zhou, Z.; Qiu, L.; Wang, L.; Zhang, H.; Gao, Y.; Wang, X.; Zhang, L.; Zhao, J.; et al. A primitive Toll-like receptor signaling pathway in mollusk Zhikong scallop *Chlamys farreri*. *Dev. Comp. Immunol.* **2011**, *35*, 511–520. [[CrossRef](#)] [[PubMed](#)]
54. Wang, M.; Wang, L.; Jia, Z.; Yi, Q.; Song, L. The various components implied the diversified Toll-like receptor (TLR) signaling pathway in mollusk *Chlamys farreri*. *Fish Shellfish Immun.* **2018**, *74*, 205–212. [[CrossRef](#)] [[PubMed](#)]
55. Wang, M.; Wang, L.; Xin, L.; Wang, X.; Wang, L.; Xu, J.; Jia, Z.; Yue, F.; Wang, H.; Song, L. Two novel LRR-only proteins in *Chlamys farreri*: Similar in structure, yet different in expression profile and pattern recognition. *Dev. Comp. Immunol.* **2016**, *59*, 99–109. [[CrossRef](#)] [[PubMed](#)]
56. Wang, M.; Wang, L.; Jia, Z.; Wang, X.; Yi, Q.; Zhao, L.; Song, L. The versatile functions of LRR-only proteins in mollusk *Chlamys farreri*. *Dev. Comp. Immunol.* **2017**, *77*, 188–199. [[CrossRef](#)] [[PubMed](#)]
57. Sharon, N. Lectins: Carbohydrate-specific Reagents and Biological Recognition Molecules. *J. Biol. Chem.* **2007**, *282*, 2753–2764. [[CrossRef](#)] [[PubMed](#)]
58. Zelensky, A.N.; Gready, J.E. The C-type lectin-like domain superfamily. *FEBS J.* **2005**, *272*, 6179–6217. [[CrossRef](#)] [[PubMed](#)]
59. Wang, L.; Wang, L.; Huang, M.; Zhang, H. The immune role of C-type lectins in molluscs. *ISJ-Invert. Surviv. J.* **2011**, *8*, 241–246.



60. Weis, W.I.; Taylor, M.E.; Drickamer, K. The C-type lectin superfamily in the immune system. *Immunol. Rev.* **1998**, *163*, 19–34. [[CrossRef](#)] [[PubMed](#)]
61. Mu, C.; Chen, L.; Zhao, J.; Wang, C. Molecular cloning and expression of a C-type lectin gene from *Venerupis philippinarum*. *Mol. Biol. Rep.* **2014**, *41*, 139–144. [[CrossRef](#)] [[PubMed](#)]
62. Wang, N.; Whang, I.; Lee, J. A novel C-type lectin from abalone, *Haliotis discus discus*, agglutinates *Vibrio alginolyticus*. *Dev. Comp. Immunol.* **2008**, *32*, 1034–1040. [[CrossRef](#)] [[PubMed](#)]
63. Zhang, J.; Qiu, R.; Hu, Y.-H. HdhCTL1 is a novel C-type lectin of abalone *Haliotis discus hannai* that agglutinates Gram-negative bacterial pathogens. *Fish Shellfish Immun.* **2014**, *41*, 466–472. [[CrossRef](#)] [[PubMed](#)]
64. Zhao, L.-L.; Wang, Y.-Q.; Dai, Y.-J.; Zhao, L.-J.; Qin, Q.; Lin, L.; Ren, Q.; Lan, J.-F. A novel C-type lectin with four CRDs is involved in the regulation of antimicrobial peptide gene expression in *Hyriopsis cumingii*. *Fish Shellfish Immun.* **2016**, *55*, 339–347. [[CrossRef](#)] [[PubMed](#)]
65. Lin, J.-Y.; Ma, K.-Y.; Bai, Z.-Y.; Li, J.-L. Molecular cloning and characterization of perlucin from the freshwater pearl mussel, *Hyriopsis cumingii*. *Gene* **2013**, *526*, 210–216. [[CrossRef](#)] [[PubMed](#)]
66. Kim, J.Y.; Adhya, M.; Cho, S.K.; Choi, K.S.; Cho, M. Characterization, tissue expression, and immunohistochemical localization of MCL3, a C-type lectin produced by *Perkinsus olseni*-infected Manila clams (*Ruditapes philippinarum*). *Fish Shellfish Immun.* **2008**, *25*, 598–603. [[CrossRef](#)] [[PubMed](#)]
67. Takahashi, K.G.; Kuroda, T.; Muroga, K. Purification and antibacterial characterization of a novel isoform of the Manila clam lectin (MCL-4) from the plasma of the Manila clam, *Ruditapes philippinarum*. *Comp. Biochem. Phys. B* **2008**, *150*, 45–52. [[CrossRef](#)] [[PubMed](#)]
68. Kang, Y.-S.; Kim, Y.-M.; Park, K.-I.; Kim Cho, S.; Choi, K.-S.; Cho, M. Analysis of EST and lectin expressions in hemocytes of Manila clams (*Ruditapes philippinarum*) (Bivalvia: Mollusca) infected with *Perkinsus olseni*. *Dev. Comp. Immunol.* **2006**, *30*, 1119–1131. [[CrossRef](#)] [[PubMed](#)]
69. Wang, L.; Song, X.; Song, L. The oyster immunity. *Dev. Comp. Immunol.* **2018**, *80*, 99–118. [[CrossRef](#)] [[PubMed](#)]
70. Zhang, H.; Wang, H.; Wang, L.; Song, L.; Song, X.; Zhao, J.; Li, L.; Qiu, L. Cflec-4, a multidomain C-type lectin involved in immune defense of Zhikong scallop *Chlamys farreri*. *Dev. Comp. Immunol.* **2009**, *33*, 780–788. [[CrossRef](#)] [[PubMed](#)]
71. Yang, J.; Huang, M.; Zhang, H.; Wang, L.; Wang, H.; Wang, L.; Qiu, L.; Song, L. CfLec-3 from scallop: An entrance to non-self recognition mechanism of invertebrate C-type lectin. *Sci. Rep.* **2015**, *5*, 10068. [[CrossRef](#)] [[PubMed](#)]
72. Wang, L.; Wang, L.; Yang, J.; Zhang, H.; Huang, M.; Kong, P.; Zhou, Z.; Song, L. A multi-CRD C-type lectin with broad recognition spectrum and cellular adhesion from *Argopecten irradians*. *Dev. Comp. Immunol.* **2012**, *36*, 591–601. [[CrossRef](#)] [[PubMed](#)]
73. Drickamer, K. Engineering galactose-binding activity into a C-type mannose-binding protein. *Nature* **1992**, *360*, 183–186. [[CrossRef](#)] [[PubMed](#)]
74. Espinosa, E.P.; Perrigault, M.; Allam, B. Identification and molecular characterization of a mucosal lectin (MeML) from the blue mussel *Mytilus edulis* and its potential role in particle capture. *Comp. Biochem. Phys. A* **2010**, *156*, 495–501. [[CrossRef](#)] [[PubMed](#)]
75. Talaei-Khozani, T.; Monsefi, M.; Ghasemi, M. Lectins influence chondrogenesis and osteogenesis in limb bud mesenchymal cells. *Glycoconj. J.* **2011**, *28*, 89–98. [[CrossRef](#)] [[PubMed](#)]
76. Zheng, P.; Wang, H.; Zhao, J.; Song, L.; Qiu, L.; Dong, C.; Wang, B.; Gai, Y.; Mu, C.; Li, C.; et al. A lectin (CfLec-2) aggregating *Staphylococcus haemolyticus* from scallop *Chlamys farreri*. *Fish Shellfish Immun.* **2008**, *24*, 286–293. [[CrossRef](#)] [[PubMed](#)]
77. Zhang, H.; Wang, H.; Wang, L.; Song, X.; Zhao, J.; Qiu, L.; Li, L.; Cong, M.; Song, L. A novel C-type lectin (Cflec-3) from *Chlamys farreri* with three carbohydrate-recognition domains. *Fish Shellfish Immun.* **2009**, *26*, 707–715. [[CrossRef](#)] [[PubMed](#)]
78. Huang, M.; Wang, L.; Zhang, H.; Yang, C.; Liu, R.; Xu, J.; Jia, Z.; Song, L. The sequence variation and functional differentiation of CRDs in a scallop multiple CRDs containing lectin. *Dev. Comp. Immunol.* **2017**, *67*, 333–339. [[CrossRef](#)] [[PubMed](#)]
79. Yang, J.; Wang, L.; Zhang, H.; Qiu, L.; Wang, H.; Song, L. C-Type Lectin in *Chlamys farreri* (CfLec-1) Mediating Immune Recognition and Opsonization. *PLoS ONE* **2011**, *6*, e17089. [[CrossRef](#)] [[PubMed](#)]

80. Yang, J.; Qiu, L.; Wei, X.; Wang, L.; Wang, L.; Zhou, Z.; Zhang, H.; Liu, L.; Song, L. An ancient C-type lectin in *Chlamys farreri* (CfLec-2) that mediate pathogen recognition and cellular adhesion. *Dev. Comp. Immunol.* **2010**, *34*, 1274–1282. [[CrossRef](#)] [[PubMed](#)]
81. Kim, Y.M.; Park, K.; Choi, K.; Alvarez, R.; Cummings, R.D.; Cho, M. Lectin from the Manila Clam *Ruditapes philippinarum* Is Induced upon Infection with the Protozoan Parasite *Perkinsus olseni*. *J. Biol. Chem.* **2006**, *281*, 26854–26864. [[CrossRef](#)] [[PubMed](#)]
82. Jia, Z.; Zhang, H.; Jiang, S.; Wang, M.; Wang, L.; Song, L. Comparative study of two single CRD C-type lectins, CgClec-4 and CgClec-5, from pacific oyster *Crassostrea gigas*. *Fish Shellfish Immun.* **2016**, *59*, 220–232. [[CrossRef](#)] [[PubMed](#)]
83. Tasumi, S.; Vasta, G.R. A Galectin of Unique Domain Organization from Hemocytes of the Eastern Oyster (*Crassostrea virginica*) Is a Receptor for the Protistan Parasite *Perkinsus marinus*. *J. Immunol.* **2007**, *179*, 3086–3098. [[CrossRef](#)] [[PubMed](#)]
84. Wang, Z.; Jian, J.; Lu, Y.; Wang, B.; Wu, Z.H. A tandem-repeat galectin involved in innate immune response of the pearl oyster *Pinctada fucata*. *Mar. Genom.* **2011**, *4*, 229–236. [[CrossRef](#)] [[PubMed](#)]
85. Zhang, D.; Jiang, S.; Hu, Y.; Cui, S.; Guo, H.; Wu, K.; Li, Y.; Su, T. A multidomain galectin involved in innate immune response of pearl oyster *Pinctada fucata*. *Dev. Comp. Immunol.* **2011**, *35*, 1–6. [[CrossRef](#)] [[PubMed](#)]
86. Song, X.; Zhang, H.; Zhao, J.; Wang, L.; Qiu, L.; Mu, C.; Liu, X.; Qiu, L.; Song, L. An immune responsive multidomain galectin from bay scallop *Argopectens irradians*. *Fish Shellfish Immun.* **2010**, *28*, 326–332. [[CrossRef](#)] [[PubMed](#)]
87. Song, X.; Zhang, H.; Wang, L.; Zhao, J.; Mu, C.; Song, L.; Qiu, L.; Liu, X. A galectin with quadruple-domain from bay scallop *Argopecten irradians* is involved in innate immune response. *Dev. Comp. Immunol.* **2011**, *35*, 592–602. [[CrossRef](#)] [[PubMed](#)]
88. Kim, J.Y.; Kim, Y.M.; Cho, S.K.; Choi, K.S.; Cho, M. Noble tandem-repeat galectin of Manila clam *Ruditapes philippinarum* is induced upon infection with the protozoan parasite *Perkinsus olseni*. *Dev. Comp. Immunol.* **2008**, *32*, 1131–1141. [[CrossRef](#)] [[PubMed](#)]
89. Chen, J.; Xiao, S.; Yu, Z. F-type lectin involved in defense against bacterial infection in the pearl oyster (*Pinctada martensii*). *Fish Shellfish Immun.* **2011**, *30*, 750–754. [[CrossRef](#)] [[PubMed](#)]
90. Badariotti, F.; Kypriotou, M.; Lelong, C.; Dubos, M.; Renard, E.; Galera, P.; Favrel, P. The Phylogenetically Conserved Molluscan Chitinase-like Protein 1 (Cg-Clp1), Homologue of Human HC-gp39, Stimulates Proliferation and Regulates Synthesis of Extracellular Matrix Components of Mammalian Chondrocytes. *J. Biol. Chem.* **2006**, *281*, 29583–29596. [[CrossRef](#)] [[PubMed](#)]
91. Xiang, Z.; Qu, F.; Wang, F.; Li, J.; Zhang, Y.; Yu, Z. Characteristic and functional analysis of a ficolin-like protein from the oyster *Crassostrea hongkongensis*. *Fish Shellfish Immun.* **2014**, *40*, 514–523. [[CrossRef](#)] [[PubMed](#)]
92. García-Maldonado, E.; Cano-Sánchez, P.; Hernández-Santoyo, A. Molecular and functional characterization of a glycosylated Galactose-Binding lectin from *Mytilus californianus*. *Fish Shellfish Immun.* **2017**, *66*, 564–574. [[CrossRef](#)] [[PubMed](#)]
93. Carneiro, R.F.; Torres, R.C.; Chaves, R.P.; de Vasconcelos, M.A.; de Sousa, B.L.; Goveia, A.C.; Arruda, F.V.; Matos, M.N.; Matthews-Cascon, H.; Freire, V.N.; et al. Purification, Biochemical Characterization, and Amino Acid Sequence of a Novel Type of Lectin from *Aplysia dactylomela* Eggs with Antibacterial/Antibiofilm Potential. *Mar. Biotechnol.* **2017**, *19*, 49–64. [[CrossRef](#)] [[PubMed](#)]
94. Kishore, U.; Ghai, R.; Greenhough, T.J.; Shrive, A.K.; Bonifati, D.M.; Gadjeva, M.; Waters, P.; Kojouharova, M.S.; Chakraborty, T.; Agrawal, A. Structural and functional anatomy of the globular domain of complement protein C1q. *Immunol. Lett.* **2004**, *95*, 113–128. [[CrossRef](#)] [[PubMed](#)]
95. Gerdol, M.; Manfrin, C.; De Moro, G.; Figueras, A.; Novoa, B.; Venier, P.; Pallavicini, A. The C1q domain containing proteins of the Mediterranean mussel *Mytilus galloprovincialis*: A widespread and diverse family of immune-related molecules. *Dev. Comp. Immunol.* **2011**, *35*, 635–643. [[CrossRef](#)] [[PubMed](#)]
96. Huang, Y.; Wang, W.; Ren, Q. Identification and function of a novel C1q domain-containing (C1qDC) protein in triangle-shell pearl mussel (*Hyriopsis cumingii*). *Fish Shellfish Immun.* **2016**, *58*, 612–621. [[CrossRef](#)] [[PubMed](#)]
97. Zhang, H.; Song, L.; Li, C.; Zhao, J.; Wang, H.; Qiu, L.; Ni, D.; Zhang, Y. A novel C1q-domain-containing protein from Zhikong scallop *Chlamys farreri* with lipopolysaccharide binding activity. *Fish Shellfish Immun.* **2008**, *25*, 281–289. [[CrossRef](#)] [[PubMed](#)]

98. Wang, L.; Wang, L.; Zhang, H.; Zhou, Z.; Siva, V.; Song, L. A C1q Domain Containing Protein from Scallop *Chlamys farreri* Serving as Pattern Recognition Receptor with Heat-Aggregated IgG Binding Activity. *PLoS ONE* **2012**, *7*, e43289. [[CrossRef](#)] [[PubMed](#)]
99. Wang, L.; Wang, L.; Kong, P.; Yang, J.; Zhang, H.; Wang, M.; Zhou, Z.; Qiu, L.; Song, L. A novel C1qDC protein acting as pattern recognition receptor in scallop *Argopecten irradians*. *Fish Shellfish Immun.* **2012**, *33*, 427–435. [[CrossRef](#)] [[PubMed](#)]
100. Lv, Z.; Qiu, L.; Wang, M.; Jia, Z.; Wang, W.; Xin, L.; Liu, Z.; Wang, L.; Song, L. Comparative study of three C1q domain containing proteins from pacific oyster *Crassostrea gigas*. *Dev. Comp. Immunol.* **2018**, *78*, 42–51. [[CrossRef](#)] [[PubMed](#)]
101. Mukhopadhyay, S.; Peiser, L.; Gordon, S. Activation of murine macrophages by *Neisseria meningitidis* and IFN- $\gamma$  in vitro: Distinct roles of class A scavenger and toll-like pattern recognition receptors in selective modulation of surface phenotype. *J. Leukoc. Biol.* **2004**, *76*, 577–584. [[CrossRef](#)] [[PubMed](#)]
102. Areschoug, T.; Gordon, S. Scavenger receptors: Role in innate immunity and microbial pathogenesis. *Cell. Microbiol.* **2009**, *11*, 1160–1169. [[CrossRef](#)] [[PubMed](#)]
103. Liu, L.; Yang, J.; Qiu, L.; Wang, L.; Zhang, H.; Wang, M.; Vinu, S.S.; Song, L. A novel scavenger receptor-cysteine-rich (SRCR) domain containing scavenger receptor identified from mollusk mediated PAMP recognition and binding. *Dev. Comp. Immunol.* **2011**, *35*, 227–239. [[CrossRef](#)] [[PubMed](#)]
104. Royet, J.; Gupta, D.; Dziarski, R. Peptidoglycan recognition proteins: Modulators of the microbiome and inflammation. *Nat. Rev. Immunol.* **2011**, *11*, 837–851. [[CrossRef](#)] [[PubMed](#)]
105. Swaminathan, C.P.; Brown, P.H.; Roychowdhury, A.; Wang, Q.; Guan, R.; Silverman, N.S.; Goldman, W.E.; Boons, G.; Mariuzza, R.A. Dual strategies for peptidoglycan discrimination by peptidoglycan recognition proteins (PGRPs). *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 684–689. [[CrossRef](#)] [[PubMed](#)]
106. Su, J.; Ni, D.; Song, L.; Zhao, J.; Qiu, L. Molecular cloning and characterization of a short type peptidoglycan recognition protein (CfPGRP-S1) cDNA from Zhikong scallop *Chlamys farreri*. *Fish Shellfish Immun.* **2007**, *23*, 646–656. [[CrossRef](#)] [[PubMed](#)]
107. Yang, J.; Wang, W.; Wei, X.; Qiu, L.; Wang, L.; Zhang, H.; Song, L. Peptidoglycan recognition protein of *Chlamys farreri* (CfPGRP-S1) mediates immune defenses against bacterial infection. *Dev. Comp. Immunol.* **2010**, *34*, 1300–1307. [[CrossRef](#)] [[PubMed](#)]
108. Itoh, N.; Takahashi, K.G. A novel peptidoglycan recognition protein containing a goose-type lysozyme domain from the Pacific oyster, *Crassostrea gigas*. *Mol. Immunol.* **2009**, *46*, 1768–1774. [[CrossRef](#)] [[PubMed](#)]
109. Iizuka, M.; Nagasaki, T.; Takahashi, K.G.; Osada, M.; Itoh, N. Involvement of Pacific oyster CgPGRP-S1S in bacterial recognition, agglutination and granulocyte degranulation. *Dev. Comp. Immunol.* **2014**, *43*, 30–34. [[CrossRef](#)] [[PubMed](#)]
110. Zhang, D.; Ma, J.; Jiang, J.; Qiu, L.; Zhu, C.; Su, T.; Li, Y.; Wu, K.; Jiang, S. Molecular characterization and expression analysis of lipopolysaccharide and beta-1,3-glucan-binding protein (LGBP) from pearl oyster *Pinctada fucata*. *Mol. Biol. Rep.* **2010**, *37*, 3335–3343. [[CrossRef](#)] [[PubMed](#)]
111. Su, J.; Song, L.; Xu, W.; Wu, L.; Li, H.; Xiang, J. cDNA cloning and mRNA expression of the lipopolysaccharide and beta-1,3-glucan-binding protein gene from scallop *Chlamys farreri*. *Aquaculture* **2004**, *239*, 69–80. [[CrossRef](#)]
112. Yang, J.; Qiu, L.; Wang, L.; Wei, X.; Zhang, H.; Zhang, Y.; Liu, L.; Song, L. CfLGBP, a pattern recognition receptor in *Chlamys farreri* involved in the immune response against various bacteria. *Fish Shellfish Immun.* **2010**, *29*, 825–831. [[CrossRef](#)] [[PubMed](#)]
113. Jayaraj, S.S.; Thiagarajan, R.; Arumugam, M.; Mullainadhan, P. Isolation, purification and characterization of  $\beta$ -1,3-glucan binding protein from the plasma of marine mussel *Perna viridis*. *Fish Shellfish Immun.* **2008**, *24*, 715–725. [[CrossRef](#)] [[PubMed](#)]
114. Jiang, S.; Wang, L.; Huang, M.; Jia, Z.; Weinert, T.; Warkentin, E.; Liu, C.; Song, X.; Zhang, H.; Witt, J.; et al. DM9 Domain Containing Protein Functions As a Pattern Recognition Receptor with Broad Microbial Recognition Spectrum. *Front. Immunol.* **2017**, *8*, 1607. [[CrossRef](#)] [[PubMed](#)]
115. Unno, H.; Matsuyama, K.; Tsuji, Y.; Goda, S.; Hiemori, K.; Tateno, H.; Hirabayashi, J.; Hatakeyama, T. Identification, Characterization, and X-ray Crystallographic Analysis of a Novel Type of Mannose-Specific Lectin CGL1 from the Pacific Oyster *Crassostrea gigas*. *Sci. Rep.* **2016**, *6*, 29135. [[CrossRef](#)] [[PubMed](#)]
116. Yang, J.; Zhao, Y.; Shao, F. Non-canonical activation of inflammatory caspases by cytosolic LPS in innate immunity. *Curr. Opin. Immunol.* **2015**, *32*, 78–83. [[CrossRef](#)] [[PubMed](#)]

117. Shi, J.; Zhao, Y.; Wang, Y.; Gao, W.; Ding, J.; Li, P.; Hu, L.; Shao, F. Inflammatory caspases are innate immune receptors for intracellular LPS. *Nature* **2014**, *514*, 187–192. [[CrossRef](#)] [[PubMed](#)]
118. Xu, J.; Shuai, J.; Li, Y.; Li, M.; Qi, C.; Zhao, D.; Yang, B.; Jia, Z.; Wang, L.; Song, L. Caspase-3 serves as an intracellular immune receptor specific for lipopolysaccharide in oyster *Crassostrea gigas*. *Dev. Comp. Immunol.* **2016**, *61*, 1–12. [[CrossRef](#)] [[PubMed](#)]
119. Lu, G.; Yu, Z.; Lu, M.; Liu, D.; Wang, F.; Wu, Y.; Liu, Y.; Liu, C.; Wang, L.; Song, L. The self-activation and LPS binding activity of executioner caspase-1 in oyster *Crassostrea gigas*. *Dev. Comp. Immunol.* **2017**, *77*, 330–339. [[CrossRef](#)] [[PubMed](#)]
120. Cebo, C.; Vergoten, G.; Zanetta, J. Lectin activities of cytokines: Functions and putative carbohydrate-recognition domains. *BBA-Biomembr.* **2002**, *1572*, 422–434. [[CrossRef](#)]
121. Zanetta, J.; Vergoten, G. Lectin domains on cytokines. *Adv. Exp. Med. Biol.* **2003**, *535*, 107–124. [[PubMed](#)]
122. Zanetta, J.; Alonso, C.; Michalski, J. Interleukin 2 is a lectin that associates its receptor with the T-cell receptor complex. *Biochem. J.* **1996**, *318*, 49–53. [[CrossRef](#)] [[PubMed](#)]
123. Cebo, C.; Dambrouck, T.; Maes, E.; Laden, C.; Strecker, G.; Michalski, J.C.; Zanetta, J.P. Recombinant Human Interleukins IL-1 $\alpha$ , IL-1 $\beta$ , IL-4, IL-6, and IL-7 Show Different and Specific Calcium-independent Carbohydrate-binding Properties. *J. Biol. Chem.* **2001**, *276*, 5685–5691. [[CrossRef](#)] [[PubMed](#)]
124. Xin, L.; Zhang, H.; Zhang, R.; Li, H.; Wang, W.; Wang, L.; Wang, H.; Qiu, L.; Song, L. CgIL17-5, an ancient inflammatory cytokine in *Crassostrea gigas* exhibiting the heterogeneity functions compared with vertebrate interleukin17 molecules. *Dev. Comp. Immunol.* **2015**, *53*, 339–348. [[CrossRef](#)] [[PubMed](#)]
125. Bilej, M.; Procházková, P.; Šilerová, M.; Josková, R. Earthworm Immunity. *Adv. Exp. Med. Biol.* **2010**, *708*, 66–79. [[PubMed](#)]
126. Beschin, A.; Bilej, M.; Magez, S.; Lucas, R.; De, B.P. Functional convergence of invertebrate and vertebrate cytokine-like molecules based on a similar lectin-like activity. *Prog. Mol. Subcell. Biol.* **2004**, *34*, 145–163. [[PubMed](#)]
127. Wolf, A.J.; Reyes, C.N.; Liang, W.; Becker, C.; Shimada, K.; Wheeler, M.L.; Cho, H.C.; Popescu, N.I.; Coggeshall, K.M.; Arditì, M.; et al. Hexokinase Is an Innate Immune Receptor for the Detection of Bacterial Peptidoglycan. *Cell* **2016**, *166*, 624–636. [[CrossRef](#)] [[PubMed](#)]
128. Seki, S.M.; Gaultier, A. Exploring Non-Metabolic Functions of Glycolytic Enzymes in Immunity. *Front. Immunol.* **2017**, *8*, 1549. [[CrossRef](#)] [[PubMed](#)]
129. Lv, Z.; Qiu, L.; Wang, W.; Liu, Z.; Xue, Z.; Yu, Z.; Song, X.; Chen, H.; Wang, L.; Song, L. A GTP-dependent Phosphoenolpyruvate Carboxykinase from *Crassostrea gigas* involved in immune recognition. *Dev. Comp. Immunol.* **2017**, *77*, 318–329. [[CrossRef](#)] [[PubMed](#)]
130. Jiang, S.; Jia, Z.; Chen, H.; Wang, L.; Song, L. The modulation of haemolymph arginine kinase on the extracellular ATP induced bactericidal immune responses in the Pacific oyster *Crassostrea gigas*. *Fish Shellfish Immun.* **2016**, *54*, 282–293. [[CrossRef](#)] [[PubMed](#)]
131. Bishop, J.R.; Gagneux, P. Evolution of carbohydrate antigens—Microbial forces shaping host glycomes? *Glycobiology* **2007**, *17*, 23R–34R. [[CrossRef](#)] [[PubMed](#)]

