Insect Biochemistry and Molecular Biology 42 (2012) 71–80

Contents lists available at SciVerse ScienceDirect

Insect Biochemistry and Molecular Biology

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

An insect TEP in a crustacean is specific for cuticular tissues and involved in intestinal defense

Chenglin Wu [a,1], Chadanat Noonin [a], Pikul Jiravanichpaisal [a,b], Irene Söderhäll [a], Kenneth Söderhäll [a,*]

[a] Department of Comparative Physiology, Uppsala University, Norbyvägen 18A, SE-752 36 Uppsala, Sweden
[b] Molecular Aquatic Biology and Genetic Laboratory, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, 113 Paholyothin, Thailand Science Park, Thailand

ARTICLE INFO

Article history:
Received 22 June 2011
Received in revised form
13 October 2011
Accepted 18 October 2011

Keywords:
A2M
TEP
Freshwater crayfish
Pacifastacus leniusculus
RNAi
Methylamine

ABSTRACT

In an attempt to identify genes encoding thioester-containing proteins in the freshwater crayfish, *Pacifastacus leniusculus*, three different cDNAs were found. A phylogenetic analysis of these proteins indicates that they can be classified into two subfamilies: two alpha-2-macroglobulins (*Pl*-A2M1, *Pl*-A2M2) showing a close similarity to shrimp A2M, and one insect TEP-like protein (*Pl*-TEP). This is the first report of an insect TEP-like protein in a crustacean. Crayfish *Pl*-A2M1, *Pl*-A2M2 and *Pl*-TEP cDNAs encode proteins with 1480, 1586 or 1507 amino acids, respectively. *Pl*-A2M1, *Pl*-A2M2 and *Pl*-TEP have the basic domain structure and functionally important residues for each molecule, and their mRNA was detected in different parts of the body, suggesting that they may have different functions. *Pl*-A2M1 was mainly expressed in hemocytes and *Pl*-A2M2 was highly expressed in heart and nerve, while *Pl*-TEP was exclusively expressed in cuticular tissues such as gill and intestine. RNA interference of *Pl*-TEP in vivo resulted in that these animals were slightly less resistant when fed with the bacterium, *Pseudomonas libanensis/gessardii*. Furthermore, when TEP activity was blocked using methylamine followed by bacterial feeding, the animals were killed to a higher extent compared to a control group. Taken together, this indicates that *Pl*-TEP and/or *Pl*-A2M1, *Pl*-A2M2 may be important for the immune defense in crayfish intestine and function as a pattern recognition protein in crayfish cuticular tissues.

1. Introduction

Thioester-containing proteins comprise a family of proteins that are characterized by a unique intrachain β-cysteinyl-γ-glutamyl thioester bond. The thioester bond mediates the binding of these proteins to their targets such as invading pathogens or attacking proteases. This protein family includes the universal protease inhibitor alpha-2-macroglobulin (A2M), complement components C3, C4, C5, glycoprotein CD109, pregnancy zone protein (PZP), C3 and PZP-like A2M domain-containing 8 (CPAMD8), and the insect TEPs which is only found in some invertebrates (Blandin and Levashina, 2004; Fujito et al., 2010).

The A2M subfamily is an abundant component in plasma of arthropods and mammals (Armstrong et al., 1996). A2M genes have been cloned and characterized in different vertebrates including humans and in invertebrates such as scallop (Ma et al., 2010), horseshoe crab *Limulus polyphemus* (Enghlid et al., 1990), mud crab *Scylla serrata* (Vaseeharan et al., 2007), and different shrimp (Ma et al., 2010). The mRNA expression of A2M was shown to be up-regulated by bacteria or white spot syndrome virus (WSSV) infection (Ho et al., 2009; Ma et al., 2010; Qin et al., 2010). The A2M contains three functional domains, a bait region, a thioester domain and a receptor binding domain (RBD). The RBD is located at the C-terminal and is exposed after proteolytic activation of A2M. The bait region locating upstream of the thioester motif showed alternative splicing (Ma et al., 2010). This alternative splicing also exists in insect TEPs, and may serve to extend the repertoire of inhibited proteases (Blandin and Levashina, 2004). *Drosophila melanogaster* TEP2 has five isoforms due to the difference in the alternative splicing region, and this region of insect TEPs is similar to the

Abbreviations: A2M, alpha-2-macroglobulin; CRD, carbohydrate recognition domain; Hpt, hematopoietic tissue; LPS, lipopolysaccharide; MBL, mannose binding lectin; PGN, peptidoglycan; PRR, pattern recognition receptor; TEPs, Thioester-containing proteins; PZP, pregnancy zone protein; CPAMD8, complement 3 and PZP-like A2M domain-containing 8.

* Corresponding author. Tel.: +46 18 4712818; fax: +46 18 4716425.
E-mail address: Kenneth.Soderhall@ebc.uu.se (K. Söderhäll).
[1] Present address: Division of Clinical Immunology, Department of Laboratory Medicine, Karolinska Institute, Stockholm, Sweden.

0965-1748/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ibmb.2011.10.006

anaphylatoxin domain in complement C3 or the bait region in A2M (Lagueux et al., 2000).

C3 proteins are mainly found in vertebrates but it can also be detected in some invertebrate deuterostome (Al-Sharif et al., 1998; Kimura et al., 2009; Nonaka et al., 1999). The C3-like genes have been identified in two main groups of protostomes; the ecdysozoan (Sekiguchi et al., in press; Zhu et al., 2005), and the lophotrochozoan (Castillo et al., 2009; Prado-Alvarez et al., 2009). In the vertebrate complement activation cascades, most proteases are classified into Bf (complement factor B) family and MASP (mannose binding lectin-associated serine protease) family. C3, Bf and MASP family genes have been identified from all invertebrate deuterostomes, but only C3 and Bf were identified from protostomes (Cerenius et al., 2010). No complement component genes have been found in insects, but 6 and 19 TEP homologus of insect TEPs are present in *D. melanogaster* and *Anopheles gambiae*, respectively. Some of them have been demonstrated to function in the immune defense as pattern recognition receptors (PRRs) (Christophides et al., 2002; Lagueux et al., 2000; Levashina et al., 2001).

The mosquito *A. gambiae* TEP1 bears some structural and functional similarities to the mammalian complement factor C3. The function of this protein is dependent on the internal thioester bond, which is demonstrated by the chemical inactivation with methylamine and by double stranded RNA knockout (Levashina et al., 2001). Mosquito TEP1 may function as a determinant of vectorial capacity in the malaria vector *A. gambiae* since its expression is up-regulated after parasite infection, and RNAi of TEP1 in adult mosquito results in higher oocyst numbers in their midguts (Blandin et al., 2004). Moreover, this protein has been reported together with two leucine-rich repeat proteins (LRRs) as major factors that regulate the load of plasmodium parasite (Fraiture et al., 2009). In *D. melanogaster*, expression of TEP1, 2 and 4 has been shown to be JAK/STAT dependent and can be up-regulated in the fat body upon bacterial challenge. This protein family has also been reported to distinguish different pathogens for subsequent phagocytosis (Lagueux et al., 2000; Stroschein-Stevenson et al., 2006). However, Bou Aoun et al. (2011) showed that *Drosophila* TEPs (TEP 1, 2, 4) are not strictly required in the body cavity to fight against bacterial and fungal infections, and therefore, these controversial results encourage further studies to clarify the role of TEPs in *D. melanogaster*, *A. gambiae* and other invertebrates.

Here, we report the cDNA cloning and characterization of three thioester-containing proteins (TEPs) from the freshwater crayfish, *P. leniusculus*. Two of them are typical A2Ms and one is an insect TEP-like protein. One of the crayfish A2Ms (*Pl*-A2M2) has previously been purified and partially sequenced (Hergenhahn et al., 1987; Hall et al., 1989). The structure of *Pl*-TEP is similar to insect TEPs, and the biological function of *Pl*-TEP was determined.

2. Materials and methods

2.1. Animals

Freshwater crayfish, *Pacifastacus leniusculus*, purchased from lake Vättern and Hjälmaren, Sweden, were kept in aquaria tap water at 10 °C. Only healthy and intermolt animals were used in the experiments.

2.2. Complementary-DNA cloning of crayfish Pl-A2M1, Pl-A2M2 and Pl-TEP

Total RNA was extracted from hemocytes, hepatopancreas, intestine, and gill using GenElute™ Mammalian Total RNA Miniprep Kit (Sigma) followed by RNase-free DNase I (Ambion) treatment. cDNA was synthesized with ThermoScript (Invitrogen) and used as a template for PCR-amplification of the TEP genes. Two pairs of degenerate primers were designed based on the amino acid sequences of the thioester region, which are highly conserved among the C3, C4 and A2M proteins of various species. The sense strand primers were CCCTCGGGNTGYGGNGARCARAAYATG and CCCTCGGGTTGYGGNGARCARANNATG (where N, R and Y represent a mix of G, A, T and C, a mix of G and A, or a mix of T and C, respectively), which correspond to the amino acid sequence PSGCGEQ (N/I/T) M at the thioester site. The antisense primers were CTTGACCACRAANGCNGTNAGCCANGT and GACCTTGACCACRWANGCNGTNAGCCA (where W represents a mix of A and T), which correspond to the amino acid sequences at about 60 amino acid residues from the C-terminal side of the thioester site (TWLTAFVVK and WLTA(F/Y)VVKV, respectively). The cDNA synthesized from crayfish total RNA of different tissues was used as a template for PCR which performed with the following condition; with 30 cycles of amplification (denaturing at 95 °C for 30 s, annealing at 55 °C for 1 min and extension at 72 °C for 1 min). A second PCR was performed using the same conditions with the first PCR products as templates. The bands of the expected size (about 230 bp) were gel-purified and cloned into the pCR2.1-TOPO vector using the TOPO-TA cloning kit (Invitrogen). The total of 60 clones was isolated and subjected to sequencing.

After verification of the DNA sequences, specific primers were designed for 5′-RACE and 3′-RACE (Table S1). Full length cDNA were indirectly obtained by using SMART RACE Kit (Clontech), in accordance with the manufacturer's instructions. Briefly, cDNA for both 5′-RACE and 3′-RACE were synthesized using ThermoScript reverse transcriptase (Invitrogen) and SMARTII A oligo (5′-AAGCAGTGGTATCAACGCAGAGTACGCGGG-3′), 5′-CDS primer A ([5′-(T)25V N-3′ (N = A, C, G, or T; V = A, G, or C)]), or 3′-CDS primer A (5′-AAGCAGTGGTATCAACGCAGAGTAC[T]30V N-3′ [N = A, C, G, or T; V = A, G, or C]). One microgram DNase I treated RNA was used in a total-volume of 10 μl reverse transcription reaction. The reverse transcription reaction was diluted with 100 μl Tricine-EDTA buffer. RACE-PCR amplification was then performed by using universal primer A (long UP: 5′-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3′; and short UP: 5′-CTAATACGACTCACTATAGGGC-3′) and GSPs. Five microliter diluted cDNA was used in 50 μl of PCR reaction with Advantage™ 2 PCR Enzyme (Clontech). The PCR program was five cycles of 94 °C for 30 s and 72 °C for 5 min; then five cycles of 94 °C for 30 s, 70 °C for 30 s and 72 °C for 5 min; followed by 25 cycles of 94 °C for 30 s and 68 °C for 30 s; and concluding with an extension cycle of 72 °C for 5 min. Amplified fragments were cloned into PCR-XL-TOPO cloning vector (Invitrogen) and subjected to DNA sequencing.

2.3. Pl-A2M1, Pl-A2M2 and Pl-TEP sequence analysis

The cDNA sequence and deduced amino acid sequence of *Pl*-A2M1, *Pl*-A2M2 and *Pl*-TEP were analyzed using the BLAST algorithm (http://www.ncbi.nlm.nih.gov/blast) and the Expert Protein Analysis System (http://www.expasy.org/). The protein domain was predicted with the simple modular architecture research tool (SMART) version 4.0 (http://www.smart.emblheidelberg.de/). Signal peptide of crayfish TEPs was predicted with SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/). The amino acid sequences of different species were obtained from NCBI database and a multiple sequence alignment was created with software ClustalX 1.83. In addition, the evolutionary relationship between crayfish TEPs representative C3, A2Ms and TEPs was investigated. Seventy one different sequences of TEPs (Table S2) were analyzed with multiple alignments and phylogenetic trees were built up with the "Neighbour-Joining (NJ)", "Maximum Parsimony (MP)", "Maximum Likelihood (ML)" or "Bayesian Interference (BI)" program. The NJ, MP and ML analyses were carried out using Mega 5 software.