Immunobiology 221 (2016) 953–963

Contents lists available at ScienceDirect

Immunobiology

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

Characterization of the gene encoding component C3 of the complement system from the spider *Loxosceles laeta* venom glands: Phylogenetic implications

D.T. Myamoto[a], G. Pidde-Queiroz[a], A. Pedroso[a], R.M. Gonçalves-de-Andrade[a], C.W. van den Berg[b], D.V. Tambourgi[a,*]

[a] Immunochemistry Laboratory, Butantan Institute, São Paulo, Brazil
[b] Institute of Molecular and Experimental Medicine, School of Medicine, Cardiff University, Cardiff, UK

ARTICLE INFO

Article history:
Received 10 March 2016
Received in revised form 28 April 2016
Accepted 23 May 2016
Available online 24 May 2016

Keywords:
Complement system
Loxosceles
Spider
C3
Gene
Phylogenetic analysis

ABSTRACT

A transcriptome analysis of the venom glands of the spider *Loxosceles laeta*, performed by our group, in a previous study (Fernandes-Pedrosa et al., 2008), revealed a transcript with a sequence similar to the human complement component C3. Here we present the analysis of this transcript. cDNA fragments encoding the C3 homologue (Lox-C3) were amplified from total RNA isolated from the venom glands of *L. laeta* by RACE-PCR. Lox-C3 is a 5178 bps cDNA sequence encoding a 190 kDa protein, with a domain configuration similar to human C3. Multiple alignments of C3-like proteins revealed two processing sites, suggesting that Lox-C3 is composed of three chains. Furthermore, the amino acids consensus sequences for the thioester was found, in addition to putative sequences responsible for FB binding. The phylogenetic analysis showed that Lox-C3 belongs to the same group as two C3 isoforms from the spider *Hasarius adansoni* (Family Salcitidae), showing 53% homology with these. This is the first characterization of a *Loxosceles* cDNA sequence encoding a human C3 homologue, and this finding, together with our previous finding of the expression of a FB-like molecule, suggests that this spider species also has a complement system. This work will help to improve our understanding of the innate immune system in these spiders and the ancestral structure of C3.

1. Introduction

Loxosceles spiders are distributed worldwide in both temperate and tropical regions and are considered as one of the most important groups of spiders in terms of public health, due to the severity of the action of their venoms (Futrell, 1992; Hogan et al., 2004). *Loxosceles* envenomation in human results in a dermonecrotic lesion and less often leads to severe systemic illness including hemolysis and renal failure (Futrell, 1992; Tambourgi et al., 2010). The main component responsible for the venom's pathophysiology is sphingomyelinase D (SMase D), which facilitates activation of the human complement system (Tambourgi et al., 1995, 1998, 2000). In order to verify the relative abundance of gene expression of SMase D, we analyzed the transcriptome of *Loxosceles laeta* venom gland from randomly selected clones from cDNA library derived through single sequencing reactions generating fragments of approximately 200–800 base pairs called expressed sequence tags (EST). ESTs represent the expressed portion of a genome and they are an important tool for purposes of gene identification and verification of gene predictions (Parkinson and Blaxter, 2009). We found that SMase D is the most expressed sequence in the transcriptome of the *Loxosceles laeta* venom glands (Fernandes-Pedrosa et al., 2008). In this study, we also found EST sequences similar to complement components C3 and factor B.

The immune system protects the organism through two layered defense mechanisms of increasing specificity, the innate and the adaptive immune responses. The innate is the oldest and is found in all multicellular organisms, while the adaptive immune system, which is highly specific and is triggered by the innate immune sys-

Abbreviations: C3, complement C3; FB, factor B; A2M, alpha-2 macroglobulin; TEP, thioester-containing proteins; TED, thioester domain; ANA, anaphylatoxin; MG, macroglobulin; C345c, C-terminal of C3, C4 and C5; FH, factor H; CR1, complement receptor 1; CR2, complement receptor 2; CR3, complement receptor 3; ORF, open reading frame; EST, expressed sequence tag; MUSCLE, Multiple Sequence Comparison by Log-Expectation; ML, Maximum Likelihood.

* Corresponding author at: Immunochemistry Laboratory, Butantan Institute, Av. Vital Brasil, 1500, Butantã, São Paulo-SP, Brazil.
E-mail address: denise.tambourgi@butantan.gov.br (D.V. Tambourgi).

http://dx.doi.org/10.1016/j.imbio.2016.05.009
0171-2985/

tem, emerged about 450 million years ago and is present in all vertebrates, except the Agnatha (Zarkadis et al., 2001).

The complement system in humans plays an essential role in both the innate and the adaptive immunity and can be activated by three pathways: classical, lectin and alternative, all converging to the cleavage of the main complement component C3, by the C3 convertases. C3 is cleaved in two fragments: C3a, a smaller fragment that recruits inflammatory cells to infection sites and C3b, a larger fragment that covalently binds to pathogen surfaces through a thioester motif, opsonizing them, promoting phagocytosis and killing by innate immune cells (Walport, 2001).

The presence of an intrachain thioester bond is a common feature of almost all proteins belonging to TEP (thioester-containing proteins) superfamily and phylogenetic studies have indicated the presence of two families: C3 and A2M (alpha-2 macroglobulin). Both families have the same basic structure of domains characterized by the presence of eight macroglobulin domains. The C3-family comprises human C3, C4 and C5 and their orthologs and differs from other TEP proteins due to the presence of ANA (anaphylatoxin) and C345c (C-terminal of C3, C4 and C5) domains, absent in the A2M family, which is composed of human A2M, pregnancy zone protein (PZP), CD109, CPDAM8 and the insect TEP orthologues (Nonaka, 2014; Sekiguchi et al., 2012). C3, C4 and C5 have probably evolved through gene duplications, however while C3 and C4 retained the intramolecular thioester bond, it has lost in C5.

The component called Cobra Venom Factor (CVF), isolated from the snake venom *Naja naja,* causes complement consumption in human and mammalian serum (Alper and Balavitch, 1976). CVF is a three-chain glycoprotein, an analog of human C3b, that binds to Factor B (FB) in human serum to generate the complex CVFB, which is cleaved by factor D (FD) resulting in a complex CVFBb, an efficient C3 and C5 convertase (Vogel et al., 1984). Unlike C3bBb and C4bC2a convertases, CVFBb is a more stable complex, since its structure, an intermediary between C3b and C3, is not made up of the TED domain and has a CUB domain in a similar position to that seen in C3b (Krishnan et al., 2009).

Comparing the three pathways, the alternative and lectin pathways seem to be the phylogenetically oldest complement activation pathways comparing to the classical pathway, which has probably developed after gene duplication events that happened in C3/C4/C5, FB/C2 and MASP/C1r/C1s and before the emergence of jawed vertebrates about 600 million years ago (Nonaka and Kimura, 2006). It is not clear yet which pathway (AP or LP) came first and there is evidence of the presence of MBL and MASP-like genes in cnidarians, ascidian and amphioxus, however there is much more evidence about the expression of the main components C3 and FB in many invertebrate species ranging from cnidarians to ascidians (Doods, 2002; Nonaka, 2011). C3- and FB-like genes were first identified in sea urchins (Al-Sharif et al., 1998; Smith et al., 1998), ascidians (Marino et al., 2002; Yoshizaki et al., 2005) and amphioxus (Suzuki et al., 2002). However, in the same period, analysis of the genomes of the nematode *C. elegans* (The C. elegans Sequencing Consortium, 1999) and of the insect *Drosophila melanogaster* (Adams et al., 2000) did not reveal any complement genes in either of these species, suggesting that the complement system has arisen in the deuterostome lineage. However, complement genes were found in limulus (Zhu et al., 2005; Ariki et al., 2008) and in cnidarians (Dishaw et al., 2005; Miller et al., 2007; Kimura et al., 2009; Fujito et al., 2010) indicating the presence of a complement system in the protostome lineage. The absence of these genes in *D. melanogaster* and *C. elegans* could be explained as a secondary loss (Nonaka and Kimura, 2006). The new interpretation of the origin of the complement system is supported by further studies that confirmed the existence of complement genes in mollusks (Zhang et al., 2007; Prado-Alvarez et al., 2009; Castillo et al., 2009; Xu et al., 2012) and other species belonging to the Phylum Arthropoda including spiders, ticks and centipedes (Sekiguchi et al., 2012; Buresova et al., 2011; Sekiguchi and Nonaka, 2015). In other arthropod groups, such as insects and diplopods, no complement genes were found nor in hydra (cnidarian), which has only an A2M gene. This suggests that the loss of C3-like proteins has occurred multiple times during the evolution of cnidarians and protostomes (Sekiguchi and Nonaka, 2015; Nonaka, 2014).

One of the ESTs, we previously found in the transcriptome of *Loxosceles laeta,* showed similarities with C3-like molecules. In order to verify if the central complement protein was expressed in venom gland, we cloned and characterized C3 from *L. laeta* venom glands, named here Lox-C3, and phylogenetically analyzed its deduced amino acid sequence.

2. Material and methods

2.1. *RNA isolation from* L. laeta *venom gland*

Loxosceles laeta spiders were collected in Campo Alegre, Santa Catarina, Brazil and kept at the Immunochemistry Laboratory of Butantan Institute, São Paulo, Brazil. Eighty *L. laeta* female spiders were subjected to food restriction to stimulate the production of mRNA in the venom glands. After 5 days, the venom glands were collected and frozen at −80 °C until use. For total RNA extraction, Trizol reagent was used following the manufacturer's instructions (Gibco-BRL Life Technologies, MD, USA). The authorization to collect *L. laeta* (permission no. 01/2009) was provided by the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA), an enforcement agency of the Brazilian Ministry of the Environment.

2.2. *RT-PCR and rapid amplification of cDNA ends (RACE)*

Based on the EST sequence LLAE0374S, which we previously identified in the transcriptome of *L. laeta* as similar to complement component C3, specific sense and antisense primers were designed to amplify the complete gene sequence (C3 forward – 5′ CGTCTTGCTTCAAGATGACC 3′ and C3 reverse—5′ GTTCCGCACTCTTGACAAGC 3′). cDNA synthesis from total RNA was performed using 5 μL (388.3 ng/μL) of total RNA and ProtoScript® M-MuLV First Strand cDNA Synthesis kit (New England Biolabs, MA, USA), following the manufacturer's instructions. Then, in order to amplify the C3-like 5′ region of *Loxosceles laeta*, two degenerate primers, comprising the TED and MG6 domain, were designed from more conserved regions based on the alignment of invertebrate C3-like sequences using BioEdit software version 7.09 (Hall, 1999) (Fig. 1). Two PCR reactions were performed. The first one using primers C3_TED forward 5′ AGYACNTGGCTNACNGCWTTY 3′ and C3 reverse 5′ GTTCCGCACTCTTGACAAGC 3′ and the second one using C3 MG6 forward 5′ GAYYTNGGHTGYGGHCCWGGH 3′ and C3_TED2 forward 5′ GAAGACGTGATGTGCAGTGGTGTG 3′. PCR amplification was performed as follows: 95 °C for 5 min, 30 cycles of 98 °C for 20 s, 63 °C for 30 s and 72 °C for 1 min and 30 s, followed by final extension at 72 °C for 5 min (Kappabiosystems, MA, USA).

Due to the difficulty in designing degenerate primers for 5′ and 3′ ends of C3-like from *L. laeta*, RACE-PCR was performed using 2.75 μL of total RNA (388.3 ng/μL) and the SMARTer RACE cDNA Amplification kit (Clontech, CA, USA), following the manufacturer's instructions to obtain 3′ and 5′ ready cDNA. To amplify the 5′ cDNA Lox-C3 gene of, we used the reverse primer ANATO – 5′ GTGCTCTGGGAATGAATCACGATCG 3′ (Integrated DNA Technologies, IA, USA) and to obtain the 3′ cDNA end, three forward primers were used: primer ANATO – 5′ GGTCCTGGTGGAGGTTTGACAAC 3′ (Síntese Biotecnologia, MG, BR), primer TED 2 – 5′ GAAGACGTGATGTGCAGTGGTGTG 3′ (Invitrogen, CA, USA) and primer C3 forward