Developmental and Comparative Immunology 53 (2015) 328–338

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

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

Identification of C-type lectin-domain proteins (CTLDPs) in silkworm *Bombyx mori*

Xiang-Jun Rao [a, 1], Toufeeq Shahzad [a, 1], Su Liu [a], Peng Wu [a], Yan-Ting He [a], Wei-Jia Sun [a], Xiang-Yun Fan [a], Yun-Fan Yang [a], Qiao Shi [a], Xiao-Qiang Yu [b, *]

[a] School of Plant Protection, Anhui Agricultural University, Hefei, Anhui 230036, PR China
[b] Division of Molecular Biology and Biochemistry, School of Biological Sciences, University of Missouri-Kansas City, Kansas City, MO 64110, USA

ARTICLE INFO

Article history:
Received 10 May 2015
Received in revised form
2 July 2015
Accepted 3 July 2015
Available online 14 July 2015

Keywords:
Insect immunity
C-type lectin
Carbohydrate recognition domain

ABSTRACT

C-type lectins (CTLs) represent a large family of proteins that can bind carbohydrate moieties normally in a calcium-dependent manner. CTLs play important roles in mediating cell adhesion and the recognition of pathogens in the immune system. In the present study, we have identified 23 CTL genes in domestic silkworm *Bombyx mori*. CTL-domain proteins (CTLDPs) are classified into three groups based on the number of carbohydrate-recognition domains (CRDs) and the domain architectures. These include twelve CTL-S (Single-CRD), six immulectins (Dual-CRD) and five CTL-X (CRD with other domains). We studied their phylogenetic features, analyzed the conserved residues, predicted tertiary structures, and examined the tissue expression profile and immune inducibility. Through bioinformatics analysis, we have putatively identified ten secretory and two cytoplasmic CTL-S; four secretory and two cytoplasmic immulectins; one secretory, one cytoplasmic and three transmembrane forms of CTL-X. Most *B. mori* CTLDPs form monophyletic groups with orthologs from Lepidoptera, Diptera, Coleoptera and Hymenoptera species. Immulectins of *B. mori* and *Manduca sexta* evolved from common ancestor genes perhaps due to gene duplication events of CTL-S ancestor genes. Homology modeling revealed that the overall structures of *B. mori* CTL domains are analogous to those of humans with a variable loop region. We examined the expression profile of CTLDP genes in naïve and immune-stimulated tissues. The expression and induction of CTLDP genes were related to the tissues and microorganisms. Together, our gene identification, sequence comparison, phylogenetic analysis, homology modeling and expression analysis laid a good foundation for the further studies of *B. mori* CTLDPs and comparative genomics.

1. Introduction

Insects use germ-line encoded pattern recognition receptors (PRRs) to recognize various pathogen-associated molecular patterns (PAMPs) and generate immune responses (Charroux et al., 2009; Ferrandon et al., 2007; Lemaitre and Hoffmann, 2007). Microorganisms carry special components on their surface that are exploited by the host immune system to distinguish between self and non-self (El Chamy et al., 2008; Kumar et al., 2009; Lemaitre et al., 1997). Toll-like receptor 4 (TLR4) mediates the immune responses to lipopolysaccharide (LPS) in humans (Poltorak et al., 1998). In *Drosophila melanogaster*, The Toll and the Imd pathway mediates the recognition of Gram (+) and Gram (−) bacteria by recognizing Lysine-type and DAP-type peptidoglycans, respectively (Lemaitre et al., 1995, 1996).

C-type lectins (CTLs) are a large family of carbohydrate-recognition proteins that exist in plants, invertebrates and vertebrates (Weis et al., 1998). CTLs usually recognize carbohydrates in a Ca^{2+}-dependent manner (Dodd and Drickamer, 2001). CTLs are involved in cell adhesion, pathogen neutralization and tailoring immune responses (Drummond and Brown, 2013; Geijtenbeek and Gringhuis, 2009; Weis et al., 1998). CTLs are a subgroup of proteins containing C-type lectin-like domain (CTLDs) or carbohydrate-recognition domains (CRDs). We collectively named them C-type lectin-domain proteins (CTLDPs) (Rao et al., 2015). CRDs with the Glu-Pro-Asn (EPN) motif are characteristic of mannose-type sugar

Abbreviations: CTL, C-type lectin; CTLD, C-type lectin-domain; CTLDP, CTL-domain protein; CRD, carbohydrate recognition domain; IML, immulectin.

* Corresponding author.
E-mail address: Yux@umkc.edu (X.-Q. Yu).

[1] These authors have made equal contribution to this study.

http://dx.doi.org/10.1016/j.dci.2015.07.005
0145-305X/

binding; CRDs with the Gln-Pro-Asp (QPD) motif are characteristic of galactose-type sugar binding (Drickamer, 1992). For human Mincle (macrophage inducible Ca^{2+}-dependent C-type lectin), E169, N171, N172, N193 and D194 are involved in calcium binding. E169 and N171 belong to the EPN motif (Furukawa et al., 2013).

We previously identified 34 CTLDP genes in the tobacco hornworm *Manduca sexta* and analyzed their sequences, structures and expression patterns. These include nine CTL-S, nineteen immulectins and six CTL-X (Rao et al., 2015). Functions of *M. sexta* IML-1 ~ 4 have been characterized. IML-1 can agglutinate bacteria and yeast in a calcium-dependent manner (Yu et al., 1999). CRD2 of IML-2 can bind to *Caenorhabditis elegans* and enhance its encapsulation and melanization *in vivo*. IML-2 is also involved in the clearance of *Serratia marcescens* and increasing the survival rates (Yu and Kanost, 2003). IML-3 contains an NWGV motif that is similar to the BH1 motif (NWGR) of mammalian galectin-3. IML-3 was observed to locate in the nucleus and stimulate hemocyte proliferation (Ling et al., 2008). IML-4 can bind to LPS, LTA (calcium-independent), *Escherichia coli*, *Staphylococcus aureus* and *Saccharomyces cerevisiae* (calcium-dependent) (Yu et al., 2006).

The silkworm *Bombyx mori* is another important lepidopteran model. Significant progress has been made in the silkworm research because of the accomplishment of genome sequencing (Xia et al., 2014, 2004). Silkworm is the major source of natural silk fibers and an economically valuable insect worldwide (Omenetto and Kaplan, 2010). Sericulture is threatened by various entomopathogens. The quantity and quality of cocoons and raw silk are affected by the frequent occurrence of silkworm diseases. The larvae of numerous lepidopterans destroy plants and cause significant economic losses. To protect sericulture and develop biological control measures against pests, it is crucial to understand the immune system of lepidopteran model insects (He et al., 2015; Jiang et al., 2010; Taniai et al., 2006; Yu et al., 2002).

Functions of several *B. mori* CTLs have been characterized (Table 1). BmLBP is produced by the fat body and binds to rough forms of bacteria and Lipid A (Koizumi et al., 1999a, 1999b, 1997). BmMBP is produced in multiple tissues and binds to different microorganisms (Watanabe et al., 2006). BmIML is inducible and produced in several tissues (Kim et al., 2003). BmLEL-1 ~ 3 are mainly produced in testis and ovary (Takase et al., 2009).

To gain an overview of *B. mori* CTLDPs, we analyzed the silkworm genome database and classified all CTLDPs into three groups: CTL-S, Immulectin (IML) and CTL-X. Orthologs were identified from other insects by sequences alignment and phylogenetic analysis. Evolutionary relationships of *B. mori* and *M. sexta* CTL domains were also analyzed. The conserved CRD residues were compared with two human CTLs. The structures of CRDs were predicted by homology modeling. The tissue profile and induced production were examined.

2. Materials and methods

2.1. Gene identification and feature analysis

CTLDP sequences from *Manduca sexta* were used as queries to search GenBank (http://www.ncbi.nlm.nih.gov/genbank/) and the Silkworm Genome Database (SilkDB, http://www.silkdb.org/silkdb/) (Duan et al., 2010; Rao et al., 2015; Wang et al., 2005; Xia et al., 2004). The cDNA and protein sequences were retrieved and confirmed in both databases. Domains and transmembrane regions were predicted in SMART (http://smart.embl-heidelberg.de/) (Schultz et al., 1998). The domain architectures were plotted with IBS 1.0 (http://ibs.biocuckoo.org/). Signal peptides were predicted with SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) (Petersen et al., 2011).

2.2. Sequence alignments and phylogenetic relationships

Multiple sequence alignments of CTLs were made with the MUSCLE unit of MEGA 6.0 at the following settings: Refining Alignment, Gap Open = −2.9, Gap Extend = 0, Hydrophobicity Multiplier = 1.2, Max Iterations = 100, Clustering Method = UPGMB, Min Diag Length = 24. The alignments were used to construct Neighbor-joining trees with the Bootstrap method, 1000 Bootstrap Replications, Substitutions Type = Amino acid, Poisson model, Uniform rates, Complete deletion of Gaps/Missing Data Treatment (Tamura et al., 2013).

2.3. Homology modeling of protein tertiary structures

Amino acid sequences of *B. mori* CTLDPs were submitted to the I-TASSER (Iterative Threading ASSEmbly Refinement) server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) for predicting and refining 3D structures (Zhang, 2008). Structural templates were first identified from PDB by the multi-threading program LOMETS (Wu and Zhang, 2007). Full-length models were then constructed by iterative simulations. The generated PDB files were rendered and visualized using the PyMOL Molecular Graphics System, Version 0.96 Schrödinger, LLC.

2.4. Tissue profile and induction analysis

Silkworm (p50T, 'Daizo' strain) eggs were kindly given by Professor Guoqing Wei in the School of Life Sciences, Anhui Agricultural University. Silkworms were reared on fresh mulberry leaves. To determine tissue expression profile of *B. mori* CTLDPs, day 3 fifth instar naïve larvae were injected with Saline, heat-killed *E. coli* strain DH5-Alpha (5×10^7 cells/larva) or *S. aureus* (5×10^7 cells/larva). Twelve hours post injection, the fat body, hemocytes, the midgut and Malpighian tubes were collected and washed in 1×

Table 1
Functions of characterized C-type lectins in *Bombyx mori*.

Name[a]	Name[b]	Inducibility	Tissue	Ligands	Referenc
BmLBP	IML-3	No	Fat body	*E. coli* (Rough) *S. Minnesota* (Rough) LPS(Lipid A)	(Koizumi et al., 1999a; Koizumi et al., 1999b; Koizumi et al., 1997)
BmMBP	IML-1	No	Hemocyte, Fat Body, etc.	*M. luteus*,*S. cerevisiae*, *C. albicans*, *S. ludwigii*, *etc.*	(Watanabe et al., 2006)
BmIML	IML-6	Yes	Fat body, Ovary, etc.	–	(Kim et al., 2003)
BmLEL-1	IML-5	Yes	Ovary, Testis,etc.	Gram-(Rough)	(Takase et al., 2009)
BmLEL-2	IML-4	Yes	Testis	Gram-(Smooth)	(Takase et al., 2009)
BmLEL-3	CTL-S4	–	Ovary, Testis,etc.	–	(Takase et al., 2009)

[a] Name in cited references.
[s] [b]Name in this article. *M. luteus*, *Micrococcu luteus*; *E. coli*, *Escherichia coli*; *S. Minnesota*, *Salmonella Minnesota*; *S. ludwigii*, *Saccharomycodes ludwigii*; *S. cerevisiae*, *Saccharomyces cerevisiae*; *C. albicans*, *Candida albicans*; LPS, lipopolysaccharide.