Available online at www.sciencedirect.com
ScienceDirect
Immunobiology 214 (2009) 165–178
Immunobiology
www.elsevier.de/imbio

Multi-component complement system of Cnidaria: *C3*, *Bf*, and *MASP* genes expressed in the endodermal tissues of a sea anemone, *Nematostella vectensis*

Ayuko Kimura, Eri Sakaguchi, Masaru Nonaka*

Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Received 5 January 2009; accepted 7 January 2009

Abstract

The origin of the complement system, one of the major systems of mammalian innate immunity, is more ancient than that of the adaptive immune system, as shown by the identification of the gene for the complement component 3 (C3) in a basic metazoa, cnidarian coral. Only a few reports on the other complement genes of non-chordates have been published, and the composition of the ancient complement system has not been clarified. We performed comprehensive cloning of the complement genes with characteristic domain structures using a Cnidarian, the sea anemone, *Nematostella vectensis*. Partial sequences of the two *C3*, two factor B (*Bf*), and one mannan-binding protein-associated serine protease (*MASP*) genes were identified in the draft genome data, and the complete coding sequences of these genes were elucidated by RT-PCR and 5′- and 3′-RACE. In contrast, no *C6* and *factor I* family genes were identified. These cnidarian components shared the unique domain structures and most of the functionally critical amino acid residues with their mammalian counterparts, suggesting the conservation of their basic biochemical functions throughout metazoan evolution. *In situ* hybridization analysis indicated that all five genes are expressed in the tentacles, pharynx, and mesentery in an endoderm-specific manner. These results suggest that the multi-component complement system comprising at least C3, Bf, and MASP was established in a common ancestor of Cnidaria and Bilateria more than 600 million years ago to protect the coelenteron, the primitive gut cavity with putative circulatory functions.

Keywords: C3; Cnidaria; Complement; Evolution; Factor B; Innate immunity; MASP

Introduction

The mammalian complement system is a highly sophisticated defense system comprising more than 30 components that are present mainly in serum and cell membranes (Volanakis 1998). Activation of this system occurs by a series of limited proteolytic cascades through three parallel pathways called the classical, alternative, and lectin pathways. These pathways merge

Abbreviations: C3, complement component 3; Bf, factor B; MASP, mannan-binding protein-associated serine protease; If, factor I; TEP, thioester bond-containing protein; A2M, alpha-2-macroglobulin; SP, serine protease; MAC, membrane attack complex; ISH, *in situ* hybridization; WISH, whole mount ISH; NJ, neighbor-joining.

*Corresponding author. Tel.: +81 3 5841 7589; fax: +81 3 5841 4990.
E-mail address: mnonaka@biol.s.u-tokyo.ac.jp (M. Nonaka).

0171-2985/$ - see front matter © 2009 Elsevier GmbH. All rights reserved.
doi:10.1016/j.imbio.2009.01.003

at the proteolytic activation step of the complement component 3 (C3), and activation fragments of C3 induce various effecter mechanisms for pathogen clearance.

Many key components of the human complement system possess unique domain structures and are classified into five mosaic protein families (Nonaka and Kimura 2006) (Fig. 1): C3 family (C3, C4, and C5); factor B (Bf) family (Bf and C2); mannan-binding protein-associated serine protease (MASP) family (MASP-1, MASP-2, MASP-3, C1r, and C1s); C6 family (C6, C7, C8A, C8B, and C9); and factor I (If) family (If only). The C3 family is included in the thioester bond-containing protein (TEP) superfamily together with the non-complement TEPs (e.g., alpha-2-macroblobulin (A2M) and CD109), although there are several C3 family-specific structural characteristics available for their identification, including the C-terminal of C3, C4, and C5 (C345c) domain at its C-terminal end. Limited proteolysis of C3 family components generates two active fragments. In the promotion of phagocytosis or induction of the cytolytic pathway, the larger fragment covalently tags the pathogens using the intramolecular thioester bond, except for C5, which secondarily lost the thioester bond. In contrast, the smaller fragment acts as an anaphylatoxin and promotes inflammation. The Bf, MASP, and If family components have a serine protease (SP) domain with trypsin-like proteolytic activity and an N-terminal region with domain combinations unique to each family that are considered responsible for their substrate specificity. The Bf and MASP family components work in the proteolytic activation cascades leading to the formation of the C3/C5 convertase, whereas If works in the proteolytic inactivation of the C3/C5 convertase. The C6 family components with the larger activation fragment of C5 are assembled into the membrane attack complex (MAC) on the membrane of pathogens that leads to the induction of cytolysis.

These complement families are thought to have become established by two steps: exon shuffling, which created the unique domain structures of each family, and gene duplication and subsequent functional divergence, which increased the number of members in each family. Accumulating information on the complement system of jawed and jawless vertebrates (Kimura et al. 2009; Nonaka and Kimura 2006) indicates that these gene duplications, which have played a pivotal role in establishing the classical pathway and the lytic pathway, occurred in the jawed vertebrate lineage. In contrast, information on invertebrate complement genes is limited, and so far only Urochordata ascidian and Cephalochordata amphioxus have been shown to have an almost complete set of the complement gene families, the *C3*, *Bf*, *MASP*, and *C6* genes (Azumi et al. 2003; Endo et al. 2003; He et al. 2008; Suzuki et al. 2002). Only a few reports have focused on the complement genes of non-chordate. So far, only the *C3* and *Bf* genes have been identified in the Echinodermata sea urchin

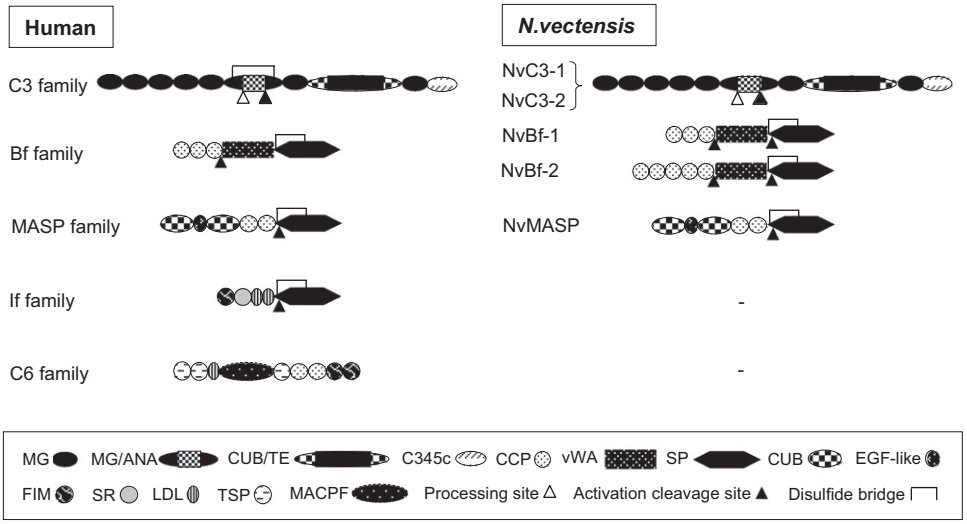

Fig. 1. Schematic drawing of the unique domain structures of the five complement families of human and *N. vectensis*. Domain structures of the five complement families of human (left) and *N. vectensis* (right) were compared. For the human C6 family, only the domain structure of C6 is shown, since the C7, C8A, C8B, and C9 lack some of the domains found in C6. Abbreviations of domain names are: MG, macroglobulin; MG/ANA, MG domain inserted with anaphylatoxin region; CUB/TE, CUB domain inserted with thioester region; C345c, C-terminal of C3, C4, and C5; CCP, complement control protein; vWA, von Willebrand factor type A; SP, serine protease; CUB, C1r, C1s, uEGF, and bone morphogenetic protein; EGF-like, epidermal growth factor-like; FIM, factor I/MAC; SR, scavenger receptor Cys-rich; LDL, low-density lipoprotein receptor domain class A; TSP, thrombospondin type 1 repeats; and MACPF, MAC/perforin. The processing site (open triangle), activation cleavage site (closed triangle), and the disulfide bridge linking two subunit chains generated by processing (line) are also indicated.