Available online at www.sciencedirect.com
ScienceDirect
Immunobiology 212 (2007) 267–277
Immunobiology
www.elsevier.de/imbio

Serine proteases of the classical and lectin pathways: Similarities and differences

Péter Gál*, László Barna, Andrea Kocsis, Péter Závodszky

Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, Karolina u. 29, Budapest H-1113, Hungary

Received 25 July 2006; received in revised form 2 November 2006; accepted 7 November 2006

Abstract

C1r, C1s, MBL-associated serine protease (MASP)-1, MASP-2 and MASP-3 are mosaic serine proteases of the classical and lectin pathways of complement. They form a family of enzymes with identical domain organization and similar overall structure, but with different enzymatic properties. MASP-2 of the lectin pathway can autoactivate and cleave C4 and C2 components. In the classical pathway two enzymes mediate these functions: C1r autoactivates and activates C1s, while C1s cleaves C4 and C2. The substrate specificity and the biological function of MASP-1 and MASP-3 have not yet been completely resolved. MASP-1 can autoactivate and the activated MASP-1 has more relaxed substrate specificity than the other members of the family. It was demonstrated that MASP-1 can specifically cleave C2, C3 and fibrinogen, but the physiological relevance of these findings has to be proved. We do not know how MASP-3 becomes activated and its biological function is also not clear. In this review, we will summarize current knowledge about the structure and function of these proteases. Special emphasis will be laid on the specificity, autoactivation and evolution of these enzymes.

Keywords: Autoactivation; Complement system; 3D structure; Evolutionary markers; Innate immunity; Serine protease; Substrate specificity

Introduction

The complement system represents one of the major effector arms of the immune response. As a part of the innate immunity it provides the first line of defence against invading pathogens, but it plays an important role in developing and modulating the adaptive immune response, as well. The complement system comprises more than 35 proteins including soluble plasma proteins and cell-surface-bound components. Of major significance in the system are serine proteases (SPs) that activate each other in a cascade-like manner. According to our present knowledge there are three different pathways through which the complement system can be activated: the classical, lectin and alternative pathways. The classical and lectin pathways have several characteristics in common, which distinguish them from the alternative pathway. The activation of the classical and lectin pathways results in the formation of the

Abbreviations: BoroMpg, Z-D-Phe-Pro-methoxypropylboroglycine-pinanediol ester; CCP, complement control protein; CUB, C1r/C1s, sea urchin uEGF and bone morphogenetic protein-1; EGF, epidermal growth factor; IGFBP, insulin-like growth factor-binding protein; MAp, MBL-associated protein; MASP, MBL-associated serine protease; MBL, mannose-binding lectin; SP, serine protease; tPA, tissue-type plasminogen activator

*Corresponding author. Tel.: +36 1 279 3135; fax: +36 1 466 5465.
E-mail address: gal@enzim.hu (P. Gál).

0171-2985/$ - see front matter © 2006 Elsevier GmbH. All rights reserved.
doi:10.1016/j.imbio.2006.11.002

C4b2a enzyme complex, which is the C3-convertase. These activation pathways are initiated when a recognition molecule binds to the appropriate activation structure. Upon this binding SP zymogens, which are associated with the recognition molecules, become activated and cleave the next components of the cascade. The classical pathway has been studied for a long time and, although there are several open questions, there is a consensus between researchers regarding the major characteristics of the pathway (Arlaud et al., 1987, 2001; Schumaker et al., 1987). The research of the lectin pathway is a relatively new area of molecular immunology, since it was discovered some 20 years ago (Ikeda et al., 1987). However despite the similarity with the well-characterized classical pathway and despite intensive research there are many disputed issues regarding the structure, function and physiological relevance of the individual components and the entire pathway, as well.

The recognition molecule of the classical pathway is C1q. C1q is composed of 18 polypeptide chains (6A, 6B and 6C) and has six globular heads and six collagen-like stalks. The globular heads bind to the antibody components of the immune complexes, whereas the collageneous stalks harbour the tetramer of C1r and C1s SPs. In the C1s–C1r–C1r–C1s tetramer both proteases have their own distinct functions: C1r is capable of autoactivation and cleavage of zymogen C1s, while activated C1s can cleave C4 and C2, the next components of the cascade.

In the case of the lectin pathway, we have many types of recognition molecules including mannose-binding lectin (MBL) (Holmskov et al., 2003) and ficolins (H, L and M-types) (Matsushita and Fujita, 2001). These molecules resemble C1q except that they are composed of one type of polypeptide chain and they bind to pathogen-associated molecular patterns (PAMPs) on the surface of microbes. Moreover MBL and ficolins exist in different oligomeric forms which have different affinity for pathogens, different associated protease composition and different complement fixation ability. The picture becomes even more complicated if we take into account the different MBL-associated serine proteases (MASPs). The most abundant protease components of the lectin pathway are MASP-1 (Matsushita and Fujita, 1992) (6 μg/ml in serum) and MASP-2 (Thiel et al., 1997) (0.5 μg/ml) while MASP-3 (Dahl et al., 2001) is regarded as a minor component. There is a consensus among researchers that MASP-2 can autoactivate and specifically cleave C4 and C2, but the substrate specificity and physiological role of MASP-1 and MASP-3 are rather debated issues. There is another MBL-associated protein: MAp19 or sMAp (Stover et al., 1999; Takahashi et al., 1999) which is the N-terminal third of MASP-2 without any catalytic activity. Combining the different recognition molecules with the different MASPs using different stoichiometry, we can get an extremely complex inventory of supramolecular complexes with potentially diverse biological functions. The complexity of the lectin pathway could be the reason for the different and often contradicting research reports published in the literature.

In this review, we summarize some aspects of our current knowledge about the C1r, C1s and MASPs of the classical and lectin pathways. During recent years, much information has accumulated about the structure and functional properties of these enzymes. These novel findings are leading to a more detailed understanding of the activation and control of the complement system, as well as providing insight into the structural background of the action of modular SPs.

The structure of C1r, C1s and MASPs

C1r, C1s, MASP-1, MASP-2 and MASP-3 form a family of SPs with identical domain organization and similar overall structure. All these enzymes are multidomain SPs consisting of six domains (Fig. 1). The C-terminal SP domain is preceded by five non-catalytic modules. The N-terminal CUB domain is followed by an epidermal growth factor (EGF)-like module, a second CUB domain and a tandem repeat of complement control protein (CCP) modules. The chymotrypsin-like SP domain (S1 peptidase family) has its N-terminal activation peptide, which is often called the ''linker region'' followed by the SP domain. These enzymes are synthesized as one-chain proenzymes and become activated when an Arg–Ile bond is cleaved between the activation peptide and the SP domain. The resulting two chains (A chain and B chain) are held together by a disulphide bridge. The A chains have a rather similar length (approximately 420–440 amino acids), while the B chain of MASP-3 (280 amino acids) is somewhat longer than the SP domains of the other four proteases (242 amino acids for C1r and MASP-2, and 251 amino acids for C1s and MASP-1). The MASP-2 protein has no post-translational modification while the A chains of the other four proteases are N-glycosylated (two glycosylation sites for C1r and C1s, and four glycosylation sites for MASP-1/3). In the case of C1r and MASP-3 the SP domain is also glycosylated (two and four glycosylation sites, respectively). It should be noted that the MASP-1 and MASP-3 proteins are the products of alternative splicing from a single *MASP-1/3* gene (Dahl et al., 2001). They share the same N-terminal CUB1–EGF–CUB2–CCP1–CCP2 region, while the SP domains are encoded by different exons. The 171-amino acid-long MAp19 protein is the alternatively spliced product of the *MASP-2* gene. It comprises the CUB1 and EGF modules of MASP-2 plus an extra C-terminal EQSL tetrapeptide, which is encoded by a separate