



HAL
open science

Structures of the MASP Proteases and Comparison with Complement C1r and C1s

Christine Gaboriaud, Véronique Rossi, Nicole Thielens

► **To cite this version:**

Christine Gaboriaud, Véronique Rossi, Nicole Thielens. Structures of the MASP Proteases and Comparison with Complement C1r and C1s. Uday Kishore; Taruna Madan; Robert B. Sim. The Collectin Protein Family and Its Multiple Biological Activities, Springer International Publishing, pp.73-101, 2021, 10.1007/978-3-030-67048-1_4. hal-03240171

HAL Id: hal-03240171

<https://hal.univ-grenoble-alpes.fr/hal-03240171>

Submitted on 28 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Structures of the MASP proteases and comparison with complement C1r and C1s

Christine Gaboriaud, Véronique Rossi and Nicole M. Thielens

University Grenoble Alpes, CEA, CNRS, IBS, Grenoble, France

Corresponding author: Christine Gaboriaud (Christine.gaboriaud@ibs.fr)

Several recognition proteins of the defence collagen family associate with proteases to initiate the complement cascade. The associated proteases, which are the subject of this review, mediate the proteolytic activation trigger. MBL-associated serine proteases (MASPs) mainly activate the lectin complement pathway (LP), while C1r and C1s activate the classical complement pathway (CP). This chapter will briefly introduce the current structural knowledge on these effector proteases and question what we know on MASPs structures, their common properties and how they differ from the C1r/s proteases. We will primarily focus on the structural comparison of the N-terminal domains, where the collectin binding sites are located and highlight novel aspects on their interaction with collagens. We also aim to highlight further molecular details associated to functional specificities, and to mention questions remaining open and needing further investigations. This chapter complements other reviews that describe the main general lines of complement activation mechanisms (1), the proposed structure-based scheme of the lectin pathway activation (2) and previous comparisons of LP and CP proteases (3,4).

1. Common protease modular structure and mode of interaction with defence collagens

MASP-1, MASP-2, MASP-3, C1r and C1s serine proteases share identical mosaic organization (Figure 1). The enzymatic activity of their C-terminal serine protease (SP) domain is controlled by the five preceding modules: two CUB (for complement C1r/C1s, Uegf, Bmp1) modules separated by an epidermal growth factor (EGF)-like module, and a pair of complement control protein (CCP) modules. LP and CP proteases are synthesized as single polypeptide chain proenzymes and become activated when a specific Arg-Ile bond is cleaved after the activation peptide in the SP domain, resulting in two chains, A and B, held together by a disulphide bridge (Figure 1). As will be detailed later, the activation cleavage is mainly triggered when the associated defence collagen binds to a target surface.

Alternative splicing products also modulate LP activity. MASP-1 and MASP-3 are the products of a single *MASP1* gene. They contain identical A chains (except for the 15 C-terminal

residues), while the B chains are encoded by different exons. In addition, alternative splicing of the *MASP1* and *MASP2* genes produces two non-enzymic proteins, called MAP-1 (also named MAp44) and MAP-2 (also named MAp19 or sMAP) (see Figure 1 and refer to (5) for the latest update on complement nomenclature). MASP-1, MASP-3, C1r and C1s have four, seven, four and two glycosylation sites, respectively, whereas MASP-2 is not glycosylated (Figure 1).

There are no X-ray crystal structures available for the full-length proteases, because of inter-domain flexibility. A dissection strategy has been setup early to unveil structure/function details of the CP and LP proteases (6–13). Two structurally independent functional regions have been identified in these complement proteases, an N-terminal interaction region (CUB1-EGF-CUB2) and a catalytic region comprising two CCP modules and the serine protease domain (CCP1-CCP2-SP). Therefore, the structural descriptions and comparisons in this review will mainly focus on these two independent functional domains. The X-ray crystal structures of several fragments within the interaction and catalytic regions have been solved for the LP and CP proteases (Figure 1).

LP proteases are probably less flexible than CP proteases, and they stably associate with the prototype MBL collectin tetramer. The overall scheme of MBL/MASPs association, illustrated on Figure 2A, and the main structural principles of the activation of the LP have been reviewed in details (2). As a main principle, MASPs associate as homodimers through the interaction domain (in green), which also mediates their association with various collectins (MBL, ficolins, CL-L1 (CL-10), CL-K1 (CL-11) or CL-LK). The protease binding sites are at the same position in the collectin collagen stems (orange), thus lying in a plane parallel to the carbohydrate surface recognized by the collectins. The height of this plane will be specific and constant for each defined collectin type. The catalytic domains (pink and grey) protrude out of the collagen cone in roughly the same plane, if one considers that the modular junctions are essentially rigid, as shown in Figure 2A. Investigations of the MBL/MASP-1 complex, combining small-angle X-ray scattering (SAXS) and electron microscopy (EM) studies, further provided experimental observations in line with such a global scheme (14). As revealed recently (15), similar interaction principles are observed for the interaction of the C1r and C1s proteases with C1q within the C1 complex (Figure 2B), although the position of the C1s catalytic domain differs.

As shown in Figure 2, the interaction regions mediate protease homodimerization (MASPs) or heterodimerization (C1r/C1s). These regions are also involved in the protease interaction with the collagen-like region of collectins (LP proteases) or of C1q (CP proteases).

2. Central CUB1-EGF interactions in LP and CP proteases assembly: a marked similarity

Early biochemical analyses demonstrated that a general feature of the CUB1-EGF regions of the MASPs is their ability to form Ca²⁺-dependent homodimers (16,17), a property shared with the CUB1-EGF region of C1s (18). Although the CUB1-EGF fragment of C1r appears as a monomer when analysed by gel filtration in the presence of calcium, it associates as a Ca²⁺-dependent heterodimer with the corresponding fragment of C1s and with full-length C1s (18–20).

The crystal structures of the CUB1-EGF segment of MASP-2 (MAp19) (11) and of C1s (9) each revealed a head to tail assembly. In both cases, the linear dimer is held by interactions between the CUB1 module of one monomer and the EGF module of its counterpart. The EGF Ca²⁺-binding sites are centrally located (green circles in Figure 3A) and stabilize this interface likely explaining the Ca²⁺-dependence of the dimerization, even if the accessibility of the calcium ions to EDTA slightly differs between the different CUB1-EGF module pairs (16,20). The dimer interface is stabilized by a combination of evenly distributed hydrophobic interactions and hydrogen bonds that are mostly conserved or substituted by similar residues in the MASP/C1r/C1s family (9,11). The main structural features of the interaction domains which are common to the LP and CP proteases are illustrated in Figure 3A, with a comparative focus on the area surrounding the calcium-binding sites in Figure 3B to 3D.

The EGF-like module, with one major and one minor anti-parallel double-stranded β -sheets and three disulphide bridges, exhibits the characteristic EGF fold. As predicted from the occurrence of a consensus calcium-binding motif, each EGF module of the MASP/C1r/C1s family binds a calcium ion (Figure 3B-D). It is coordinated in a conserved manner by seven oxygen ligands, including six homologous residues contributed by the EGF modules and a water molecule that also forms an H-bond with a conserved Gly residue in loop 4 of the CUB1 module (Figure 3) (9,11). Additional CUB1-EGF interactions provide further stabilization of the intramonomer CUB1-EGF interface (dark green in Figure 3B-D). Conserved residues stabilizing the dimer interface are shown with light-green highlights and sticks (3C, 3D). A particular feature of MASP-1/3 and C1r EGF lies in the presence of a cluster of charged residues in loop Legf (Figure 3C), which is unusually large in C1r EGF and is mostly disorganized, in contrast with the corresponding loop in MASP-2 and C1s. The role of this charged, surface-exposed and flexible loop in potential interactions is currently unknown.

The CUB1 modules exhibit a typical compact ellipsoid structure. They differ from the canonical topology of spermadhesins, assembled from two five-stranded antiparallel β -sheets, by lacking the first two β -strands. Such a deletion is observed solely in the CUB1 module of the MASP/C1r/C1s family (21). Unexpectedly, unveiling the first X-ray structures human C1s and MASP-2 CUB1-EGF fragments revealed the presence of a Ca²⁺-binding site on the distal edge of the CUB1 module (9,11). Coordination of the Ca²⁺ ion was found to involve three acidic residues (red

and orange, Figure 3B-D) that, together with a closely associated tyrosine residue, are conserved in a large proportion of the CUB repertoire and define a signature of Ca-binding CUB (cbCUB) modules (21). All CUB modules of the MASP/C1r/C1s family belong to this cbCUB domain subset (Figure 3C-D). The Ca²⁺ ion is the central element of a network of interactions that stabilize the distal end of the CUB module. However this ion is exposed to the solvent and readily exchangeable, which likely explains why the Ca²⁺-binding site was partly disordered and free in the X-ray structure of the CUB1-EGF-CUB2 fragment of rat MASP-2, crystallized in the absence of calcium (10). Recently published crystal structures of the same rat fragment confirmed the presence of Ca²⁺ ions in each of the CUB modules (22).

The currently available X-ray crystal structures (Figure 1) of CUB domains, CUB1-EGF and CUB1-EGF-CUB2 fragments of human and rat MASP-1/3 and MASP-2, and of human C1r and C1s have further confirmed these similarities among the LP and CP proteases, especially for the central mode of calcium-dependent association of CUB1-EGF pairs (green circle in Figure 3A), and for the structural motifs defining the three calcium-binding sites (Figure 3C-D).

3. Similar interaction properties mediate protease association with defence collagens

The CUB1-EGF fragment of MASPs has been shown to bind to MBL and ficolins in a Ca²⁺-dependent manner. However, the affinity increases when the CUB2 module is also present, i. e. in the full-length proteases or in the CUB1-EGF-CUB2 fragments of MASPs (10⁻⁸ M versus 10⁻⁹ M range) (23,24). These observations suggested that the interaction of MASPs with MBL and ficolins involves a major contribution of the N-terminal CUB1-EGF module pair, but is strengthened by the following CUB2 module.

Mutational analyses of CUB1 and CUB2 modules identified homologous binding sites in MASP-1/3 and MASP-2 for MBL and ficolins, located in close vicinity of their Ca²⁺-binding sites (11,23), in loops L5 and L9 (yellow and orange residues in Figure 3B). On the collectin side, mutagenesis studies in the collagen-like regions of rat or human MBL and ficolins identified a conserved lysine residue essential for the interaction with the MASPs (25–28). This information allowed to propose an initial model of the complex of a tetramer of MBL and the MASP-1/3 CUB1-EGF-CUB2 dimer, featuring four major homologous interaction sites involving the conserved lysine residue of the collagen-like region of MBL and acidic Ca²⁺ ligands of CUB1 and of CUB2 of the protease (23). This fits with the more complete model later proposed in Figure 2A (2).

Whereas it was initially proposed that the interaction of the C1s-C1r-C1r-C1s tetramer with C1q differed from that of the MASP dimers with MBL or ficolins (29), two 2009 studies (17,30) provided evidence for similarities in the collagen binding sites of the initiating complexes of the CP

and LP, despite different stoichiometries. Phillips et al. demonstrated *in vitro* cross-interaction of the subcomponents of the lectin and classical pathways, although cognate interactions were tighter than non-cognate interactions, the latter being more transient with faster dissociation rates (17).

High affinity C1q binding sites in the C1r CUB1 and CUB2 modules and a lower affinity site in C1s CUB1 module were identified using site-directed mutagenesis of C1r and C1s (30). As observed previously for MASPs, these sites involve acidic residues also contributing Ca²⁺ ligands (red and orange in Figure 3C-D). A refined model of the C1 complex was built, in which the CUB1-EGF-CUB2 of C1r and the CUB1-EGF of C1s interact through six binding sites with homologous lysines of the C1q stems, nicely reconciling the different stoichiometry of C1q and of its associated proteases (30).

As shown previously in Figure 3C-D, several residues are conserved in the CUB1 and CUB2 modules of all these proteases, suggesting similarities in the interaction mechanism for the initiating complexes of LP and CP. Mutation experiments have shown that the Ca²⁺-binding site not only plays a stabilizing role of the distal end of the CUB domain, but also provides two exposed acidic residues involved in the interaction with a lysine residue of the collagen-like triple helix. Different models of C1 and MBL-MASP complexes derived from modelling and binding studies proposed analogous interactions of C1q with the C1s-C1r-C1r-C1s tetramer and of tetrameric MBL with MASP dimers, with binding sites present on each CUB domain, except for C1s CUB2 (31). Quite similarly, lysine attraction by the acidic calcium-binding residues was also observed for the binding between cobalamin-bound gastric intrinsic factor and the CUB6 module of cubilin, revealing a common mechanism for calcium-dependent CUB interactions (21,32).

However, C1s CUB2 stands out as an exception regarding this common interaction scheme. Indeed, the mutation of the homologous acid residue in C1s CUB2 did not impact C1 formation, revealing that this module was not involved in C1q binding, although its calcium binding property is intact (30). We can observe that the conserved tyrosine and histidine residues in L5 (highlighted yellow) and the conserved serine residue in L9 (orange) are missing in C1s CUB2, thus suggesting their essential role for collagen binding (Figure 3C, 3D). At this stage, we can also note that these loops L5 and L9 are significantly conserved within each CUB1 and CUB2 subsets (except for C1s CUB2), but significantly differ between CUB1 and CUB2.

4. Opposite binding orientations for CUB1 and CUB2 modules with collagen: a structural view

X-ray structural analyses rapidly confirmed the common mode of interaction suggested above (Figure 4). The first structure solved (33) showed the CUB2 domain of rat MASP-1/3 in

complex with a 27-residue synthetic collagen-like peptide containing a central lysine residue (to mimic the MASP binding site of MBL). The X-ray crystal structure of a complex of the same collagen-like peptide and the interaction region of human C1s (CUB1-EGF-CUB2 dimer) was solved later by the same team (34). The two structures confirmed the Ca^{2+} -dependency of the MASP-MBL interaction, with the side chain of a collagen lysine (K15 corresponding to Lys46 in rat MBL-A) contacting three residues involved in Ca^{2+} coordination (Figure 4A, 4B). Intriguingly however, the fine structural comparison of the complexes of this same collagen peptide with MASP-1/3 CUB2 (33) or C1s CUB1 (34) reveals a striking binding orientation difference, as illustrated on Figure 4A, 4B, and with their superposition on Figure 4E. This difference in orientation is observed despite the common interaction details described above. Does it reflect a different mode of interaction between MASPs and C1s, or between CUB1 and CUB2 modules? As mentioned before, CUB1 and CUB2 modules are conserved in the MASPs and C1r/C1s family (at the exception of C1s CUB2). But CUB1 and CUB2 readily differ in the size and composition of loops L5 and L9 (Figure 3C, 3D), and these loops are involved in collagen binding (Figure 4A, 4B). So we can hypothesize that different CUB1 and CUB2 relative orientations could be dictated by their differences in loops L5 and L9.

How details of CUB-collagen interactions explain the differences in orientation:

CUB1 interaction details (34). In this structure, the side chain of lysine 15 mainly interacting with the CUB1 module of C1s comes from the **trailing** collagen strand (purple, Figure 4C). It contacts the carboxylate groups of Glu45 and Asp98, and the hydroxyl and carbonyl groups of Ser100, three Ca^{2+} -coordinating residues of CUB1 (Figure 4C). Additional contacts involve hydrogen bonds provided by Tyr52 of C1s (L5). Glutamic acid residues 48 (L5) and 102 (L9) clamp the two sides of the collagen helix. Since glutamic acid residue 48 interacts with the lysine 15 of the collagen middle strand (grey), we can note that two consecutive collagen lysines are recognized, at least in this specific C1s example. However, this secondary interaction may not be essential, since single lysine mutations of C1q can prevent C1 assembly (30). Of note, this glutamic acid 48 residue is conserved in CP but not in LP proteases.

CUB2 interaction details (33). In that structure, the side chain of lysine 15 interacting with the CUB2 module of rat MASP-1/3 comes from the **leading** collagen strand (dark blue, Figure 4D). It interacts with the carboxylate groups of Glu216 and Asp263 and the hydroxyl group of Ser265 (orange), three Ca^{2+} -coordinating residues of CUB2. Additional interacting contributions come from residues in loop L5 (His218 and Tyr225 which occupy a groove in the collagen helix), which is stabilized by its close proximity to the collagen peptide. L5 and L9 loops thus occupy two consecutive grooves in the collagen helix.

The analysis of these two structures shed light on the role of loops L5 and L9 in collagen interaction. The residues involved in collagen binding are mostly conserved in the LP and CP

proteases (except for C1s CUB2). But, as shown previously, the loops L5 and L9 significantly differ between CUB1 and CUB2 (Figure 3). As discussed by Girija et al. (34), the loop L9 in CUB2 would prevent binding in the CUB1 orientation, because of steric clashes. Thus, this difference in binding orientations may be a general feature of CUB1/CUB2 modules, acquired through evolution. Interestingly though, this difference in orientation fits with the constraint that CUB1 and CUB2 bind to two opposite collagen stems (Figure 4A, 4B and global model in Figure 2). Moreover, as illustrated on Figure 4E, superimposition of the two X-ray structures opens the hypothesis that CUB1 and CUB2 modules could simultaneously bind to the same collagen stem. Such a possibility has not been explored in the initial models but is supported by the snapshot EM structure of the IgM/C1/C4b complex, where C1r CUB1 and CUB2 are observed to bind to the same collagen stem of C1q (Figure 4F, (15)).

5. CUB1-EGF-CUB2 homodimer (LP) versus heterodimer (CP) interactions: common features and differences

Comparing CUB-EGF-CUB sequences reveals strong similarities between LP and CP proteases. As mentioned before, CUB1-EGF head-to-tail dimer association is supported by tightly conserved sequences in the interaction domain of the LP and CP proteases. Sequence conservation further extends in the next EGF-CUB2 segment, as shown in Figure 5A. Indeed, we could build a consensus sequence motif spanning: (i) the end of the EGF module, including residues involved in the CUB1-EGF dimer interface (boxed in green), (ii) the 3 residues long linker (in grey) and (iii) the beginning of the CUB2 module (in salmon). Remarkably, scanning sequence databases with this consensus sequence enables to retrieve most LP or CP protease homologs of different species (more than 800 sequences in translated EMBL). Importantly, scanning sequences with this motif does not retrieve other unrelated EGF-CUB containing protein. This, again, stresses how these proteases are similar.

However, by comparing different CUB1-EGF-CUB2 dimer structures from LP and CP proteases, we can observe striking differences in the relative CUB2 orientation, as illustrated in Figure 5B-D. CUB1-EGF-CUB2 homodimers of LP proteases are structurally highly similar (22,23). The C1s CUB1-EGF-CUB2 homodimer exhibits more flexibility but retains a similar overall shape, in the sense that CUB2 modules are not involved in the dimer interaction (Figure 5C). In contrast, C1r-C1s heterodimers present a larger buried dimer interface, with additional contacts between CUB1 and CUB2 modules (Figure 5D), which likely explains their preferential heterodimer association (35). These differences thus seem to correlate with homo versus heterodimer (C1r/C1s) association status.

Are these conformational differences related to fine sequence specificities? In the sequence alignment (Figure 5A), the residues involved in the CUB1-EGF dimer interface (boxed in green) are far more conserved than the ones involved in the MASP EGF-CUB2 monomer interface (boxed in light blue). One ionic interaction (label [1] on Figure 5A and 5B) is highlighted, because it is conserved in MASPs EGF-CUB2 interface, but not in C1r/s. As another striking difference, the three-residues long linker sequence (grey) starts with a glycine residue in C1s only (conserved in more than 150 species), which likely enhances the flexibility of its EGF-CUB2 junction.

Therefore, in contrast to the strong stability and similarity of the CUB1-EGF junction and interface, as described above, there is a significant flexibility at the EGF-CUB2 junction in the C1s molecule, as revealed by the variable and unusual position of the C1s CUB2 (Figure 5B-C) in different X-ray structures. The recent cryo-EM study of the C1/IgM/C4b complex further confirmed this flexibility, since the C1s CUB2 module is seen there in another orientation (15,34). Although probably less flexible than its C1s counterpart, the EGF-CUB2 junction also adopts a different conformation in C1r (Figure 5D).

The above mentioned local sequence differences between MASPs and C1r/s are in line with the fact that MASPs mainly associate as homodimers whereas C1r and C1s interaction domains preferentially associate as heterodimers. In theory, heterodimer formation in MASP proteases is possible, but only via subunit exchange following EDTA dissociation. This concerns not only MASPs with identical interaction regions (MAp44/MAP-1, MASP-1 and MASP-3) (36), but also MASP-1, MASP-3 together with MASP-2 (37). Some free MASPs heterocomplexes might be found in blood, but the majority of MASPs moieties are associated to collectin recognition molecules (36). Nevertheless, the main difference in favoured associations between CP and LP proteases, the wide versatility in the collectin recognition proteins and stoichiometries in complexes with MASPs, versus the strict C1 composition, really make a difference between LP and CP proteases.

6. How CCP1 and CCP2 modulate LP and CP catalytic activities

The catalytic regions of LP and CP proteases, comprising the two CCP modules and the SP domain, mediate enzymatic activity. This activity implies a conformational transition in the C-terminal SP domain, to convert the proenzyme into an active state. As for coagulation proteases, activation of complement proteases proceeds through rapid cascade amplification with increased serum concentrations of the successive activation targets (C4/C2, C3). The activation of the LP or CP initiating complexes involves two successive steps. In the case of LP, MASP-1 auto-activates “on site”, when its associated collectin pattern recognition molecule (PRM) binds to a target surface; activated MASP-1 will then activate MASP-2. The CP pathway proceeds quite similarly, with a first step of C1r auto-activation and C1s activation by activated C1r in a second step. These activations

are triggered by C1 binding to an appropriate activating surface, the main prototype surface being IgM or hexameric IgG immune complexes platforms. For both LP and CP cascades, the next common step is C4 cleavage (by MASP-2 or C1s) and C2 cleavage (C2 being associated to C4b) to form the CP/LP convertase. One difference is that C1r and C1s are associated to the same C1q molecule, which is not the standard case for MASP-1 and MASP-2 in LP activating complexes.

From a biochemical point of view, the activation of LP or CP protease is driven by the specific proteolytic cleavage of a specific Arg-Ile bond in the N-terminal part of its SP domain, between the A and B chains, which remain attached through a disulphide bridge (see Figure 1). From a structural standpoint, this cleavage is mediated by a partner SP domain, in an overall relative orientation similar to the one shown in Figure 6A, impeded in the relative orientation shown in Figure 6B.

Several structures corresponding to this catalytic region have been solved for the different C1r/s and MASPs proteases (Figure 1). They correspond to different conformational states: active (activated/two chains) or proenzyme (zymogen). In order to get their X-ray structure, engineered point mutations have often been introduced to stabilize a specific conformational state (mutation of the active site, of the activation site, or introduction of patient mutations). In addition, stabilization was also achieved within complexes formed with inhibitors or cognate substrate (Figure 1).

CCP1-CCP2 as elongated arms or handles

Evolutionary conserved rigid interaction between the SP domain and the preceding module can be detected at the sequence level, as initially illustrated with the examples of these complement and blood coagulation proteases (38). The presence of two CCP (or sushi) modules preceding the C-terminal SP domain is a specific hallmark of the MASPs/C1r/C1s proteases, which they share with haptoglobin, although the latter has lost catalytic activity (39,40).

The X-ray structure of C1s CCP2-SP fragment (6) first revealed that the CCP2 module stands perpendicular to the protease surface. Therefore, a structural role of a rigid handle was initially inferred. Sequence conservation (38) suggested that this property would mainly be conserved in the different CP and LP proteases, which was confirmed later. Little or moderate flexibility has been observed between CCP1 and CCP2, thus providing a handle extension. For example, CCP1-CCP2 lay rigidly in MASP-1 (41) and only moderate flexibility has been observed for C1s (34). From the functional standpoint, these elongated arms correctly position the C-terminal SP domain, both for the activation process and for the catalytic cleavage. As seen on Figure 2, the position of C1s lying on C4 is quite different from that of MASP-1 in the MBL/MASP model, whereas the position of C1r looks similar to that of MASP-1. The CCP1-CCP2 arm will transmit and amplify displacements occurring in the preceding modules, in case of flexibility. This fits with the fact that C1s is highly

flexible at the EGF-CUB2 junction preceding the CCP1-CCP2 modules (Figure 4), with also limited flexibility at the CCP1-CCP2 junction (34) and CCP2-SP junction (42). The handle transmission is achieved through the CUB2-CCP1 connection, which is assumed to be rigid. Indeed, it is observed in a stable conformation in C1s, when comparing two different fragments containing this junction (34). It is also assumed to be rigid and conserved in MASP-1 and MASP-2, maintained by two ionic interactions and a central hydrophobic contact (2).

CCP1-CCP2 include a C4 binding exosite in MASP-2 and C1s

What is the structural background explaining why MASP-2 and C1s can both cleave C4, and not the other proteases? The first piece of answer resides in the observation that both CCP1-CCP2 modules of C1s and MASP-2 play a role in properly positioning C4 before and after its cleavage. These two modules can be exchanged between MASP-2 and C1s, the ones in MASP-2 providing more catalytic efficiency (43). The X-ray structure of the MASP-2/C4b complex clearly identified a negative patch at the CCP1 and CCP2 junction in MASP-2, which is involved in binding an arginine patch in the C345C domain of C4/C4b (2,42). Mutational analyses further confirmed the essential role of several exosite residues in C4 deposition, mostly in MASP-2 CCP2 (E333, P340, D365, P368). Another negative patch at the CCP1-CCP2 junction is found in C1s, with a quite similar relative position towards C4b observed in the cryo-EM C1/IgM/C4b envelope (15). Point mutations in C1s CCP1-CCP2 linker also revealed the importance of this sequence for C4 cleavage (44). This negative patch is not present in the other CP and SP proteases. An additional C4 binding exosite in the SP domain was also identified, as will be discussed later (2,4,42).

CCP modules potentially restricting spontaneous activation: a C1r exception

Among the LP and CP proteases, only the catalytic regions in C1r can form head-to-tail dimers through interactions between the CCP1 of one monomer and the SP domain of its counterpart (cf. Figure 1, figure 6B, (8,45). Thus, in absence of C1q and in presence of calcium, C1r dimers are at the centre of C1s-C1r-C1r-C1s tetramers (46, 47), this configuration being unique to the CP. As illustrated on Figure 6B, this head-to-tail association of C1r catalytic domains prevents activation of one monomer by its counterpart, and the C1r SP domains are somewhat protected from unwanted activation, especially within the C1s-C1r-C1r-C1s tetramer. More details about the activation process will be described in the next section related to SP domains. This observation that C1r CCP1-CCP2 modules likely prevent spontaneous C1r activation sounds physiologically relevant since C1q and the C1r and C1s proteases are not always secreted by the same cells and circulate at high concentrations. How C1-inhibitor, which is assumed to bind to the C1 proteases tetramer further controls this unwanted activation of the proteases is currently not fully understood.

7. Structural similarities in the C-terminal trypsin-like SP domain

Sequence and structural similarities assign the SP domain of MASPs and C1r/s to the largest S1A family of proteases, most commonly known as trypsin-like proteases (recently reviewed in (48)). Trypsin-like proteases are synthesized in a proenzyme state. They are activated on site through enzymatic cascade amplifications in the cases of complement and coagulation proteases. This involves a structural conversion which folds the active site for efficient binding and catalysis (49).

Bovine chymotrypsin led to the very first X-ray structure solved in this family (50). MASPs and C1r/s share therefore the chymo/trypsin-like fold and associated activation mechanism. In this family, the prototypic chymotrypsin serves as reference for the numbering scheme proposed by Hartley in 1970 for the SP domain (51), with the catalytic triad residues termed *His57*, *Asp102* and *Ser195*. In the simplest activation model early suggested for this proteases family, the new charged N-terminus, created by the activation cleavage, inserts into the activation pocket, attracted towards *Asp194* (as on Figure 6C, 6E), which triggers concerted surface loop rearrangements.

The trypsin-like primary specificity, which dictates a preferential cleavage after an arginine or lysine, is another common structural feature shared by all the CP and LP proteases. This cleavage specificity is related to the presence of a conserved aspartic residue at the bottom of the primary substrate-binding pocket (noted D189 on Figure 6). This residue is only properly positioned in the active conformation, after the concerted rearrangements of loops 1 to 3, which are essential in shaping the enzyme activity (Figure 6C, 6E).

As briefly mentioned before, C4 recognition by MASP-2 involves an additional exosite in the SP domain, which binds C4 sulfotyrosines (42). In this exosite, mutation of arginine residues 578 and 583 impaired C4 cleavage, further demonstrating their essential role (52). Cumulating mutations in the CCP and SP exosites leads to severe reduction in C4 cleavage efficiency, suggesting cumulative functional effects of these distal exosites (52). A similar exosite is found in C1s D loop, with residues K560 and R561 (Figure 6D-E). Interestingly, the conformational switch converting C1s zymogen into an active enzyme is required to properly position these two exosite C4 binding residues (Figure 6D-E, (53), reviewed in (4)). This remark holds true especially for arginine 583, which is displaced by 12 Å when comparing zymogen and active MASP-2 (52). However, MASP-2 possibly exhibits higher plasticity than C1s. These overall structural similarities between C1s and MASP-2 readily explain why only these two enzymes can cleave the large C4 component.

8. Structural differences shaping protease specificity, an opportunity to design specific inhibitors

Even if they share a common trypsin fold, proteases of the S1 family have evolved different catalytic roles. The fine catalytic specificities are mainly defined by structural differences in the loops surrounding the active site (as initially described/ reviewed by Perona & Craik (54)). The length and conformation of these loops (Figure 7) mainly differ in each of the LP and CP proteases, as well as their sequence details. Better understanding of how these loops shape substrate specificity is acquired through the structural analysis of protease/inhibitor complexes, because these loops also shape differential inhibitor sensitivities. For example, inhibitors produced in various organisms have been described, such as gigastasin, a leech inhibitor for MASP-1, MASP-2 and C1s, or the bacterial pan-inhibitor ecotin (inhibiting MASP-2, MASP-3 and, to a lesser extent, MASP-1) (55–57). X-ray structures of protease/inhibitor complexes have been solved in these two latter cases (Figure 1). Gigastasin forms tight contacts with the C1s SP domain. These contacts span all subsites, plus the anion-binding exosite involved in C4 binding, through two sulfotyrosines near the gigastasin C-terminal terminus (56).

Reversely, taking into account the specific environment provided by each set of surface loops has enabled the design of highly specific inhibitors for MASP-1, MASP-2 and MASP-3, which can lead to promising therapeutic specific control of the different proteases. Structures of several enzyme/inhibitor complexes have been solved (Figure 1). Structural details in these structures are often consistent with specific features of the designed inhibitors. For example, ecotin binds through hydrophobic residues in the S' sites of MASP-3, a feature that is also found in the specific TFMI-3 inhibitor (55,58).

These specific inhibitors have already provided precious tools to decipher and deconvolute the functional role(s) of the MASP-1 and MASP-2 proteases (59,60). They have also been used to demonstrate that MASP-3 is in fact the exclusive activator of pro-factor D in blood, making a strong connexion with complement alternative pathway activation (58).

MASP-1 SP domain is evolutionary distinct and supports a more versatile catalytic activity. One main structural basis supporting this difference is illustrated on Figure 7, where MASP-1 clearly features a more open binding site. In that sense, MASP-1 is more close to trypsin than to highly specific coagulation or complement proteases. MASP-1 is a very potent activator of the LP pathway, already in its zymogen form and even more in its active state (61). One peculiar feature of MASP-1 is that *Asp189*, at the bottom of the primary specificity pocket, is already engaged in a salt bridge with an Arg side chain. Also the B loop in MASP-1 is highly extended, as in thrombin, which may constrict substrate specificity at this level (reviewed in (61)).

In contrast to MASP-1, access to the substrate-binding site is highly restricted in MASP-2, C1s and C1r by the surface loops surrounding the active site (Figure 7). This explains why these proteases mainly act only on two substrates. Selective substrate recognition may also be further strengthened by a step of substrate induced fit. This has been suggested from the structural plasticity observed in MASP-2 (59), later confirmed in the case of its association to C4 (42). Analysis of the MASP-3/ecotin complex structure (4IW4, Figure 1) also suggested allosteric conformational changes associated to ecotin binding to an extended MASP-3 exosite (55).

9. Structural uncertainties beyond common principles in the LP and CP activation mechanism

The simple activation scheme described previously, initially discovered on digestive proteases, has also evolved in the S1 proteases family. In the case of complement activation by LP and CP proteases, the associated collectin PRM plays a major role in triggering their activation only on site. For example, the major transactivation mechanism of LP activation proposes that juxtaposition of MBL-MASP complexes sitting on the same carbohydrate surface (as in Figure 2A) will drive protease contacts triggering LP activation (2,62). This process relies on the hypothesis that the SP domains are rigidly handled, as discussed before. Interestingly, this mechanism explains how activation is restricted to activating surfaces and how the recognition of these surfaces by the associated collectin controls the activation process. A similar transactivation mechanism has been proposed for the CP (15,63), but uncertainties remain in this partly controversial issue. Further experimental evidences and details are needed to understand how the different activation events are topologically orchestrated in the context of various PRM-MASPs complexes or of C1 (46,61).

What about zymogen catalytic activity? MASP-1 zymogen is able to cleave MASP-1 and MASP-2, but active MASP-1 is far more potent on these activation cleavages (41,61). MASP-2 structural plasticity also led to the suggestion of a step of induced fit in its activation mechanism, with analogy to the “substrate-induced catalysis” proposed for the activation of complement factors D, B and C2 (59).

Finally, we need to stress that some (auto-)activation schemes observed *in vitro* may not occur as such *in vivo*, where they can be mitigated by a lack of proximity or by physiological inhibitors. Since they share identical primary cleavage specificity (after Lys or Arg), CP, LP and other trypsin-like proteases can be indeed controlled by the same inhibitor. For example, C1-inhibitor is a main common physiological inhibitor of CP, LP but also of contact proteases. C1-inhibitor, in a similar way as the other members of the serpin family, acts as suicide substrate for C1r, C1s, MASP-1 and MASP-2. How C1-inhibitor binds to LP and CP activating complexes mainly

remains an open question on the structural side. Some other inhibitors of the serpin family control the LP activation: anti-thrombin (for MASP-1 and MASP-2, (64)) and alpha2-macroglobulin (for MASP-1,(65)). We also lack structural data on this ground, except that different sensitivities to inhibitors are modulated by surface properties that shape the function and specificity of each enzyme, as mentioned before.

10. LP and CP protease enzymatic activities out of the complement system

MASP-1, with its wider substrate binding site, as seen before, can in fact cleave a larger range of substrates (3). Therefore, active MASP-1 has been described as a promiscuous protease, able to boost coagulation (by acting on prothrombin, fibrinogen and coagulation factor XIII), and contributing to a powerful inflammatory reaction through activation of endothelial cells and bradykinin release (by acting on protease-activated receptors (PARs) and kininogen). More details can be found in (61,66)

Several mutations of the *MASP-1* gene that specifically affect MASP-3 SP domain are associated with a rare developmental disorder described as the 3MC syndrome. This suggests that MASP-3 is involved in other non-canonical functions, independent of the complement system, which remain to be deciphered. The structural impact of one of these mutations, G666[c197]E, has been investigated in MASP-3 zymogen form. The X-ray structure (4KKD, Figure 1) provided evidence for a destabilizing effect on the active site, as expected, which further reinforced the hypothesis that an undefined MASP-3 enzymatic activity, abolished by this patient mutation, impacts the development in the 3MC syndrome (67). Other mutations in MASP-3 (68,69) or in collectins CL-K1 and CL-L1 (70) have also been related to this 3MC syndrome, but structural and functional details are still elusive. However, these observations strongly suggest a crucial role for at least some of these lectin pathway molecules in development processes (66).

On another side, mutations were recently identified in *C1R* (mostly) and *C1S* genes in patients affected by the periodontal Ehlers-Danlos syndrome (71). Looking for a possible common molecular effect associated to the different patient mutations, molecular analyses led to the hypothesis that active C1s, or an active 40 kDa C1s fragment, may be one common pathological component (72,73). Because the identified patient mutations either alter the enzyme correct folding or the association with C1q, this CP protease activity is assumed to be free from the control of C1q. Consistently, complement has been shown to play a role in the dysbiotic transformation of the periodontal microbiota and in the inflammatory process that leads to the destruction of periodontal bone (74). However, additional effects in the case of the periodontal Ehlers-Danlos syndrome are expected because the symptoms differ from classical periodontitis (71). Non-

canonical functions of these C1r and C1s proteases, out of the complement system, are likely to be involved in this and other diseases processes.

General comments and conclusion

Our current structural knowledge on the LP and CP proteases has dramatically increased during the past decades, starting from the resolution of the X-ray structure of C1s catalytic domain in 2000 (6). These homologous proteases share many structural similarities, as illustrated in Figures 2, 3, 4 and 6. However, several differences can be highlighted when we compare LP and CP proteases, or each individual protease. On one hand, MASPs mainly associate as homodimers and bind to a wide variety of collectin molecules: MBL (which shows various oligomeric assemblies), the three ficolins and collectins (CL-K1, CL-L1 and CL-LK). On the other hand, the C1 stoichiometry is fixed and the physiological concentration of its components is higher. C1r and C1s probably briefly transit as a proenzyme tetramer before their association with C1q, in a configuration that can be interpreted as a way to prevent self-activation, as described above. Such configuration is a distinctive feature of the CP proteases.

From an evolutionary point of view, C1r, C1s, MASP-2, MASP-3, as well as thrombin are more recent than MASP-1. The following two evolution marks support this idea: the AGY codon for the active serine (instead of TCN) and the absence of introns in the SP domain (3,75). MASP-1 is thus atypical, with a wider spectrum of substrates in its active state (76). Nevertheless MASP-1 activity is crucial in activating the LP (61). As briefly reminded in the last chapter section, recent studies of patient mutations in MASP-3, C1r, C1s have suggested further functional roles independent of the complement system. This opens the way to further structure/function studies associated to diseases.

Questions about the partial activity of proenzyme states as a function of protease plasticity have been addressed elsewhere (3,61,77). We would like to mention that a collapsed conformation in C1r zymogen structure further prevents substrate entry and involves another conformational transition towards the un-collapsed state (78). Another collapsed proenzyme conformation has been observed for MASP-1. MASP-2 was shown to partly cleave C4 in its proenzyme state. This is the subject of a minor controversial issue, related to the shift in the position of the two arginines of its C4 exosite in the SP domain, as shown in this review for the homologous case of C1s (Figure 6) (52). Nevertheless, such a tight control of C4 cleavage by C1s or MASP-2, which mainly requires that the proteases are activated, is in line with the strong need for complement activation only “on site”, under the control of their associated recognition molecules.

The overall flexibility of these CP and LP proteases has limited their study by X-ray analyses to recombinant fragments. This has however provided insights into the relative flexibility of joints linking two successive modules, by comparing their positions in different structures. For example, the CUB1-EGF head-to-tail associations are remarkably conserved in all these CP and LP proteases, as detailed in section 2. MASP CUB2-CCP1 junction appears to be mainly rigid and mostly conserved in MASP-1 and MASP-3 (2). This CUB2-CCP1 is also likely rigid in C1s (34). As mentioned in section 6.1, restricted flexibility at the CCP1-CCP2 junction further supports the initial notion that it can act as a handle or elongated arm. In contrast, far more flexibility has been observed at the EGF-CUB2 junction in the CP proteases, especially in C1s. Complementary studies using SAXS to analyse the soluble structure of full-length MASP-1, MASP-2 and of the C1 proteases tetramer, combined with rigid body modelling and negative stain EM allowed to propose low resolution models of the full-length protease assemblies, which mainly exhibit extended shapes (14,22,47,63). These experiments performed in solution provide however far less experimental values, which introduces a higher risk of over- or miss-interpretation. Within the current resolution revolution in cryo-EM techniques, cryo-electron tomography has brought into the field the only current mid-resolution snapshot of the full-length C1r and C1s proteases within the IgM-C1-C4b giant complex (Figure 2B, (15), which represents a fantastic advance in our insight into these molecules. Further insights on this side are awaited, that will complete our understanding of the activation scenario of the CP proteases (as compared to LP), with the topological details of successive steps. How and where serpin inhibitors such as C1-inhibitor are engaged and control some of these steps further remain to be investigated.

Funding

The work presented here was partly funded by French National Research Agency Grants ANR-09-PIRI-0021 and ANR-16-CE91-0004-01.

Figure legends

Figure 1. MASPs and C1r/s share identical mosaic structures.

(A) Modular structure of the MASPs related proteins (left) and C1r/s (right). The serine proteases are activated by a specific cleavage between the A and B chains, which are held together by a disulphide bridge. The activation peptide at the N-terminus of the Serine Protease domain is shown in white. Glycosylation sites are shown with black diamonds. B) This table lists the X-ray structure coordinates deposited in the Protein Data Bank, with their accession code and their schematic representation, such as head-to-tail dimers or stars indicating the presence of protein ligands (inhibitor in red, ligand/substrate in yellow). Enzyme active forms are indicated using a space separation between the two chains, since activation cleavage splits A and B chains. Corresponding references numbers are indicated.

Figure 2. Overall view of LP or CP activating complexes assembly.

(A) and (B) Main global model of interaction between collectins (here MBL) and MASPs proteases. The N-terminal CUB1-EGF-CUB2 interaction domains (here in green) binds inside the cone defined by the collagen stems (orange helices), with the catalytic domains pointing outside the cone (pink + grey). The orientation and height of the catalytic Serine Protease domains (SP, grey) is defined by the association of the proteases within the collectin bound to a glycan surface, held by a rigid handle composed of CUB2-CCP1-CCP2 modules. Side (A) and top (B) views. Figure extracted from the review by Kjaer et al. (2) with publisher's permission. (C) Cryo-EM snapshot of IgM/C1/C4b mega complex structure. As labelled, the C1s and C1r proteases are shown in pink and purple, respectively; C1q and C4b in dark and light blue, respectively. The protease interaction domains lie in a plane parallel to the IgM binding platform (green rectangle). Two perpendicular views are shown, which highlight the different orientations of the catalytic domains of C1r (left) and C1s (right). Figure adapted from Sharp et al (15).

Figure 3. Structural similarity of EGF, CUB1 and CUB2 calcium binding sites.

(A) Overall relative spatial orientation of the calcium-binding sites in MASP1/3 CUB1-EGF-CUB2 dimer. (B) Relative position in CUB1-EGF-CUB2 monomer (MASP-1/3) of the zones around each calcium-binding site. (C) The corresponding conserved sequence features in MASPs and C1r/s of these zones around calcium-binding sites: Ca_EGF, Ca_CUB1 and Ca_CUB2. Ca_EGF is at the boundary between CUB1 and EGF, see arrows above defining the transition between the two modules. Conserved aspartic residues with a bidentate calcium binding are in red. Other conserved residues are shown in orange (calcium ligands), light green (conserved CUB-EGF dimer interface), dark green (intramonomer CUB1-EGF interface) or yellow (binding to collectin). Marked sequence

difference in C1s CUB2 module, which does not take part directly in C1q binding, is highlighted in light pink and marked by a blue #. CUB L5 and L9 loops, essential elements of interaction with the recognition protein, are labelled. (D) Comparative views of the CUB1-EGF monomeric interface (Ca_EGF) as well as CUB1 and CUB2 calcium-binding sites in MASP1/3, C1r and C1s. Positions where conserved binding residues are missing in C1s CUB2 are displayed with pink sticks and highlighted with a blue circle. Same residue colour code as described above in (C). Figure drawn using PDB files 3DEM (MASP1/3), 6F1C (C1r and C1s CUB2), 1NZI (C1s CUB1 and EGF). Human MASP-1 is shown here as a representative example of MASPs structures which are very similar, as shown in (22).

Figure 4. Structural insight into the interaction of CUB modules with a MBL-like collagen peptide or with C1q.

(A) C1s CUB1 (PDB code 4LOR, (34)) and (B) MASP-1/3 CUB2 (PDB code 3POB, (33)) structures in complex with a MBL-like collagen peptide. The lysine residues of the common collagen ligand are shown as stick (dark blue, leading strand; grey, middle strand; purple, trailing strand). (C) and (D) Focus on interaction details of CUB1 (4LOR) and CUB2 (3POB) binding with collagen. Major residues are labelled. Same colour code as above and as in Figure 3 for the CUB residues. Note that E48 in C1s (L5), which mediates direct electrostatic interactions with the middle collagen lysine, is not conserved in MASPs sequences. E) The two structures above are shown after superimposition of the three essential lysines of the common collagen peptide (same colour code and stick representation as above in (C) and (D)). This superimposition suggests that one CUB1 and one CUB2 module can interact with the same collagen stem. (F) CUB1 and CUB2 modules from two C1r chains interact with the same C1q collagen stem in the context of the IgM/C1/C4b mega complex structure investigated by cryo-EM, as highlighted by stars (adapted from (15)).

Figure 5. Consensus and variations of the EGF-CUB2 interface in homo versus hetero-dimer contexts.

(A) Strong sequence similarity between MASPs and C1r/s proteins in a segment spanning from the end of EGF (green) to the beginning of CUB2 (salmon), the three amino-acid long linker being displayed in grey. The consensus sequence motif shown below can be used to retrieve all MASPs or C1r/s homologs in other species. Label [I] indicates a conserved electrostatic interaction in MASPs (letter coloured red for aspartic acid and blue for lysine residues). The green highlights indicate conserved CUB1-EGF dimer interface residues. The light blue highlights show residues interacting at the EGF-CUB2 interface in MASPs. The red dots above the sequence are conserved calcium ligands. Majuscule letters show positions defined in the consensus sequence. Bold letters show

significant differences between LP and C1r or C1s proteases; r_MASP2: rat MASP-2 (B) and (C) Homodimer CUB-EGF-CUB in MASP-1/3 (PDB code 3DEM, (23)) and MASP-2 (PDB code 5CIS), where the linker fragment is coloured in grey, and the two residues forming an electrostatic interaction [I] in red and blue, as in (A). The CUB1-EGF-CUB2 orientation is indicated. (D) Homodimer CUB1-EGF-CUB2 in C1s (PDB code 4LMF,(34)). Same colour code. (E) Heterodimeric C1s/C1r CUB1-EGF-CUB2 association (PDB code 6F1C, (35)). Same colour code. All dimers are shown with the same orientation of the CUB1-EGF central association.

Figure 6. SP domain activation: relative SP domain positions and activation conformational changes.

(A) Relative overall orientation required for activating cleavage, where the bond to be cleaved (activation site) is close to the active site of the activating molecule (into the red circle). The SP domain to be activated (*to be act.*) and the activating SP domain (*act.*) are labelled. This C1r structure corresponds approximately to an enzyme product complex, since Arg446 of the cleaved activation loop (*to be act.*) sits in the S1 substrate-binding pocket of the other molecule (*act.*) (PDB code: 2QY0). (B) The head-to-tail dimeric structure of C1r catalytic domains, held by contacts between the SP domain and the CCP1 module of its partner (PDB code 1GPZ, major association also observed in 2QY0). In this orientation, the active site and activation sites are 90 Å distant (red arrows), at both ends of the dimer. This distance might help preventing unwanted autoactivation cleavage in the context of the C1s-C1r-C1r-C1s tetramer (8). A pioneer electron micrograph of the tetramer is shown above, the structure shown here corresponds to the central yellow circle, in a perpendicular view (47, 8). (C) The activation conformational change is illustrated with the superimposed overview of active (yellow, 1ELV) and proenzyme (salmon, 4J1Y) states of C1s. The active serine (S195) is shown red. The common backbone is shown in grey (active) and dark (proenzyme). Two C4 exosite residues are shown in magenta. Other important residues discussed in the text are shown. (D) and (E) Focus on the active site details in the active (D) and proenzyme states (E). Same details as in (C). The aspartic acid at the bottom of the primary pocket (D189), the aspartic (D194) residue interacting with the new N-terminal extremity released by the activation cleavage (Nt) are labelled. K560 and R561 are essential residues in the C4 binding exosite (4). In (D), the green segment and arginine side-chain show the position of the target C4 fragment inside the binding cleft (following superposition with the MASP-2/C4 complex, PDB code 5JPM, (79)). The nomenclature of the loops A to E is the one defined by Perona & Craik (54), with loops 1 to 3 mostly involved in the activating conformational transition.

Figure 7. Structural comparison of surface loops around the active site.

(A-B) The active site (red) is surrounded by several loops, which define various interacting surfaces. A widely open active site was initially observed in digestive proteases, and this is only seen in MASP-1 (A), which fits with its more versatile enzymatic activity. In contrast, active site access is highly restricted in the C1s protease (B) and MASP-2, which mainly act on only two substrates, namely complement C4 and C2. More surface comparisons have been shown in (80). (C) Superimposition of LP and CP proteases showing the variation of length and conformation around the catalytic triad (orange sticks). The reference structure is the ecotin/MASP-3 complex (PDB code 4IW4), and thus the primary binding segment of ecotin (yellow) shows the position of the canonical substrate binding site. Same loop labels as in Figure 6, Panel C from Gaboriaud et al. (55).

References

1. Merle NS, Church SE, Fremeaux-Bacchi V, Roumenina LT. Complement System Part I - Molecular Mechanisms of Activation and Regulation. *Front Immunol* (2015) **6**:262. doi:10.3389/fimmu.2015.00262
2. Kjaer TR, Thiel S, Andersen GR. Toward a structure-based comprehension of the lectin pathway of complement. *Mol Immunol* (2013) **56**:413–422. doi:10.1016/j.molimm.2013.05.007
3. Gál P, Barna L, Kocsis A, Závodszky P. Serine proteases of the classical and lectin pathways: similarities and differences. *Immunobiology* (2007) **212**:267–277. doi:10.1016/j.imbio.2006.11.002
4. Pike RN, Wijeyewickrema LC. The molecular switches controlling the interaction between complement proteases of the classical and lectin pathways and their substrates. *Curr Opin Struct Biol* (2013) **23**:820–827. doi:10.1016/j.sbi.2013.07.016
5. Bohlsón SS, Garred P, Kemper C, Tenner AJ. Complement Nomenclature-Deconvoluted. *Front Immunol* (2019) **10**:1308. doi:10.3389/fimmu.2019.01308
6. Gaboriaud C, Rossi V, Bally I, Arlaud GJ, Fontecilla-Camps JC. Crystal structure of the catalytic domain of human complement c1s: a serine protease with a handle. *EMBO J* (2000) **19**:1755–1765. doi:10.1093/emboj/19.8.1755
7. Budayova-Spano M, Grabarse W, Thielens NM, Hillen H, Lacroix M, Schmidt M, Fontecilla-Camps JC, Arlaud GJ, Gaboriaud C. Monomeric structures of the zymogen and active catalytic domain of complement protease c1r: further insights into the c1 activation mechanism. *Structure* (2002) **10**:1509–1519. doi:10.1016/s0969-2126(02)00881-x
8. Budayova-Spano M, Lacroix M, Thielens NM, Arlaud GJ, Fontecilla-Camps JC, Gaboriaud C. The crystal structure of the zymogen catalytic domain of complement protease C1r reveals that a disruptive mechanical stress is required to trigger activation of the C1 complex. *EMBO J* (2002) **21**:231–239. doi:10.1093/emboj/21.3.231
9. Gregory LA, Thielens NM, Arlaud GJ, Fontecilla-Camps JC, Gaboriaud C. X-ray structure of the Ca²⁺-binding interaction domain of C1s. Insights into the assembly of the C1 complex of complement. *J Biol Chem* (2003) **278**:32157–32164. doi:10.1074/jbc.M305175200

10. Feinberg H, Uitdehaag JCM, Davies JM, Wallis R, Drickamer K, Weis WI. Crystal structure of the CUB1-EGF-CUB2 region of mannanose-binding protein associated serine protease-2. *EMBO J* (2003) **22**:2348–2359. doi:10.1093/emboj/cdg236
11. Gregory LA, Thielens NM, Matsushita M, Sorensen R, Arlaud GJ, Fontecilla-Camps JC, Gaboriaud C. The X-ray structure of human mannan-binding lectin-associated protein 19 (MAp19) and its interaction site with mannan-binding lectin and L-ficolin. *J Biol Chem* (2004) **279**:29391–29397. doi:10.1074/jbc.M402687200
12. Harmat V, Gál P, Kardos J, Szilágyi K, Ambrus G, Végh B, Náray-Szabó G, Závodszky P. The structure of MBL-associated serine protease-2 reveals that identical substrate specificities of C1s and MASP-2 are realized through different sets of enzyme-substrate interactions. *J Mol Biol* (2004) **342**:1533–1546. doi:10.1016/j.jmb.2004.07.014
13. Gál P, Harmat V, Kocsis A, Bián T, Barna L, Ambrus G, Végh B, Balczer J, Sim RB, Náray-Szabó G, et al. A true autoactivating enzyme. Structural insight into mannanose-binding lectin-associated serine protease-2 activations. *J Biol Chem* (2005) **280**:33435–33444. doi:10.1074/jbc.M506051200
14. Kjaer TR, Le LTM, Pedersen JS, Sander B, Golas MM, Jensenius JC, Andersen GR, Thiel S. Structural insights into the initiating complex of the lectin pathway of complement activation. *Structure* (2015) **23**:342–351. doi:10.1016/j.str.2014.10.024
15. Sharp TH, Boyle AL, Diebolder CA, Kros A, Koster AJ, Gros P. Insights into IgM-mediated complement activation based on in situ structures of IgM-C1-C4b. *Proc Natl Acad Sci USA* (2019) **116**:11900–11905. doi:10.1073/pnas.1901841116
16. Thielens NM, Cseh S, Thiel S, Vorup-Jensen T, Rossi V, Jensenius JC, Arlaud GJ. Interaction properties of human mannan-binding lectin (MBL)-associated serine proteases-1 and -2, MBL-associated protein 19, and MBL. *J Immunol* (2001) **166**:5068–5077. doi:10.4049/jimmunol.166.8.5068
17. Phillips AE, Toth J, Dodds AW, Girija UV, Furze CM, Pala E, Sim RB, Reid KBM, Schwaeble WJ, Schmid R, et al. Analogous interactions in initiating complexes of the classical and lectin pathways of complement. *J Immunol* (2009) **182**:7708–7717. doi:10.4049/jimmunol.0900666
18. Thielens NM, Aude CA, Lacroix MB, Gagnon J, Arlaud GJ. Ca²⁺ binding properties and Ca²⁺(+)-dependent interactions of the isolated NH₂-terminal alpha fragments of human complement proteases C1-r and C1-s. *J Biol Chem* (1990) **265**:14469–14475.
19. Busby TF, Ingham KC. Calcium-sensitive thermal transitions and domain structure of human complement subcomponent C1r. *Biochemistry* (1987) **26**:5564–5571. doi:10.1021/bi00391a052
20. Thielens NM, Enrie K, Lacroix M, Jaquinod M, Hernandez JF, Esser AF, Arlaud GJ. The N-terminal CUB-epidermal growth factor module pair of human complement protease C1r binds Ca²⁺ with high affinity and mediates Ca²⁺-dependent interaction with C1s. *J Biol Chem* (1999) **274**:9149–9159. doi:10.1074/jbc.274.14.9149
21. Gaboriaud C, Gregory-Pauron L, Teillet F, Thielens NM, Bally I, Arlaud GJ. Structure and properties of the Ca²⁺-binding CUB domain, a widespread ligand-recognition unit involved in major biological functions. *Biochem J* (2011) **439**:185–193. doi:10.1042/BJ20111027

22. Nan R, Furze CM, Wright DW, Gor J, Wallis R, Perkins SJ. Flexibility in Mannan-Binding Lectin-Associated Serine Proteases-1 and -2 Provides Insight on Lectin Pathway Activation. *Structure* (2017) **25**:364–375. doi:10.1016/j.str.2016.12.014
23. Teillet F, Gaboriaud C, Lacroix M, Martin L, Arlaud GJ, Thielens NM. Crystal structure of the CUB1-EGF-CUB2 domain of human MASP-1/3 and identification of its interaction sites with mannan-binding lectin and ficolins. *J Biol Chem* (2008) **283**:25715–25724. doi:10.1074/jbc.M803551200
24. Skjoedt M-O, Roversi P, Hummelshøj T, Palarasah Y, Rosbjerg A, Johnson S, Lea SM, Garred P. Crystal structure and functional characterization of the complement regulator mannose-binding lectin (MBL)/ficolin-associated protein-1 (MAP-1). *J Biol Chem* (2012) **287**:32913–32921. doi:10.1074/jbc.M112.386680
25. Wallis R, Shaw JM, Uitdehaag J, Chen C-B, Torgersen D, Drickamer K. Localization of the serine protease-binding sites in the collagen-like domain of mannose-binding protein: indirect effects of naturally occurring mutations on protease binding and activation. *J Biol Chem* (2004) **279**:14065–14073. doi:10.1074/jbc.M400171200
26. Teillet F, Lacroix M, Thiel S, Weilguny D, Agger T, Arlaud GJ, Thielens NM. Identification of the site of human mannan-binding lectin involved in the interaction with its partner serine proteases: the essential role of Lys55. *J Immunol* (2007) **178**:5710–5716. doi:10.4049/jimmunol.178.9.5710
27. Girija UV, Dodds AW, Roscher S, Reid KBM, Wallis R. Localization and characterization of the mannose-binding lectin (MBL)-associated-serine protease-2 binding site in rat ficolin-A: equivalent binding sites within the collagenous domains of MBLs and ficolins. *J Immunol* (2007) **179**:455–462. doi:10.4049/jimmunol.179.1.455
28. Lacroix M, Dumestre-Pérard C, Schoehn G, Houen G, Cesbron J-Y, Arlaud GJ, Thielens NM. Residue Lys57 in the collagen-like region of human L-ficolin and its counterpart Lys47 in H-ficolin play a key role in the interaction with the mannose-binding lectin-associated serine proteases and the collectin receptor calreticulin. *J Immunol* (2009) **182**:456–465. doi:10.4049/jimmunol.182.1.456
29. Gaboriaud C, Teillet F, Gregory LA, Thielens NM, Arlaud GJ. Assembly of C1 and the MBL- and ficolin-MASP complexes: structural insights. *Immunobiology* (2007) **212**:279–288. doi:10.1016/j.imbio.2006.11.007
30. Bally I, Rossi V, Lunardi T, Thielens NM, Gaboriaud C, Arlaud GJ. Identification of the C1q-binding Sites of Human C1r and C1s: a refined three-dimensional model of the C1 complex of complement. *J Biol Chem* (2009) **284**:19340–19348. doi:10.1074/jbc.M109.004473
31. Wallis R, Mitchell DA, Schmid R, Schwaeble WJ, Keeble AH. Paths reunited: Initiation of the classical and lectin pathways of complement activation. *Immunobiology* (2010) **215**:1–11. doi:10.1016/j.imbio.2009.08.006
32. Andersen CBF, Madsen M, Storm T, Moestrup SK, Andersen GR. Structural basis for receptor recognition of vitamin-B(12)-intrinsic factor complexes. *Nature* (2010) **464**:445–448. doi:10.1038/nature08874
33. Gingras AR, Girija UV, Keeble AH, Panchal R, Mitchell DA, Moody PCE, Wallis R. Structural basis of mannan-binding lectin recognition by its associated serine protease MASP-1: implications for complement activation. *Structure* (2011) **19**:1635–1643. doi:10.1016/j.str.2011.08.014

34. Venkatraman Girija U, Gingras AR, Marshall JE, Panchal R, Sheikh MA, Harper JAJ, Gál P, Schwaeble WJ, Mitchell DA, Moody PCE, et al. Structural basis of the C1q/C1s interaction and its central role in assembly of the C1 complex of complement activation. *Proc Natl Acad Sci USA* (2013) **110**:13916–13920. doi:10.1073/pnas.1311113110
35. Almitairi JOM, Venkatraman Girija U, Furze CM, Simpson-Gray X, Badakshi F, Marshall JE, Schwaeble WJ, Mitchell DA, Moody PCE, Wallis R. Structure of the C1r-C1s interaction of the C1 complex of complement activation. *Proc Natl Acad Sci USA* (2018) **115**:768–773. doi:10.1073/pnas.1718709115
36. Rosbjerg A, Munthe-Fog L, Garred P, Skjoedt M-O. Heterocomplex formation between MBL/ficolin/CL-11-associated serine protease-1 and -3 and MBL/ficolin/CL-11-associated protein-1. *J Immunol* (2014) **192**:4352–4360. doi:10.4049/jimmunol.1303263
37. Paréj K, Hermann A, Donáth N, Závodszy P, Gál P, Dobó J. Dissociation and re-association studies on the interaction domains of mannan-binding lectin (MBL)-associated serine proteases, MASP-1 and MASP-2, provide evidence for heterodimer formation. *Mol Immunol* (2014) **59**:1–9. doi:10.1016/j.molimm.2013.12.003
38. Gaboriaud C, Rossi V, Fontecilla-Camps JC, Arlaud GJ. Evolutionary conserved rigid module-domain interactions can be detected at the sequence level: the examples of complement and blood coagulation proteases. *J Mol Biol* (1998) **282**:459–470. doi:10.1006/jmbi.1998.2008
39. Tosi M, Duponchel C, Meo T, Couture-Tosi E. Complement genes C1r and C1s feature an intronless serine protease domain closely related to haptoglobin. *J Mol Biol* (1989) **208**:709–714. doi:10.1016/0022-2836(89)90161-7
40. Redmond AK, Ohta Y, Criscitiello MF, Macqueen DJ, Flajnik MF, Dooley H. Haptoglobin Is a Divergent MASP Family Member That Neofunctionalized To Recycle Hemoglobin via CD163 in Mammals. *J Immunol* (2018) **201**:2483–2491. doi:10.4049/jimmunol.1800508
41. Megyeri M, Harmat V, Major B, Végh Á, Balczer J, Héja D, Szilágyi K, Datz D, Pál G, Závodszy P, et al. Quantitative characterization of the activation steps of mannan-binding lectin (MBL)-associated serine proteases (MASPs) points to the central role of MASP-1 in the initiation of the complement lectin pathway. *J Biol Chem* (2013) **288**:8922–8934. doi:10.1074/jbc.M112.446500
42. Kidmose RT, Laursen NS, Dobó J, Kjaer TR, Sirotkina S, Yatime L, Sottrup-Jensen L, Thiel S, Gál P, Andersen GR. Structural basis for activation of the complement system by component C4 cleavage. *Proc Natl Acad Sci USA* (2012) **109**:15425–15430. doi:10.1073/pnas.1208031109
43. Rossi V, Teillet F, Thielens NM, Bally I, Arlaud GJ. Functional characterization of complement proteases C1s/mannan-binding lectin-associated serine protease-2 (MASP-2) chimeras reveals the higher C4 recognition efficacy of the MASP-2 complement control protein modules. *J Biol Chem* (2005) **280**:41811–41818. doi:10.1074/jbc.M503813200
44. Bally I, Rossi V, Thielens NM, Gaboriaud C, Arlaud GJ. Functional role of the linker between the complement control protein modules of complement protease C1s. *J Immunol* (2005) **175**:4536–4542. doi:10.4049/jimmunol.175.7.4536
45. Kardos J, Harmat V, Palló A, Barabás O, Szilágyi K, Gráf L, Náráy-Szabó G, Goto Y, Závodszy P, Gál P. Revisiting the mechanism of the autoactivation of the complement protease C1r in the C1 complex: structure of the active catalytic region of C1r. *Mol Immunol* (2008) **45**:1752–1760. doi:10.1016/j.molimm.2007.09.031

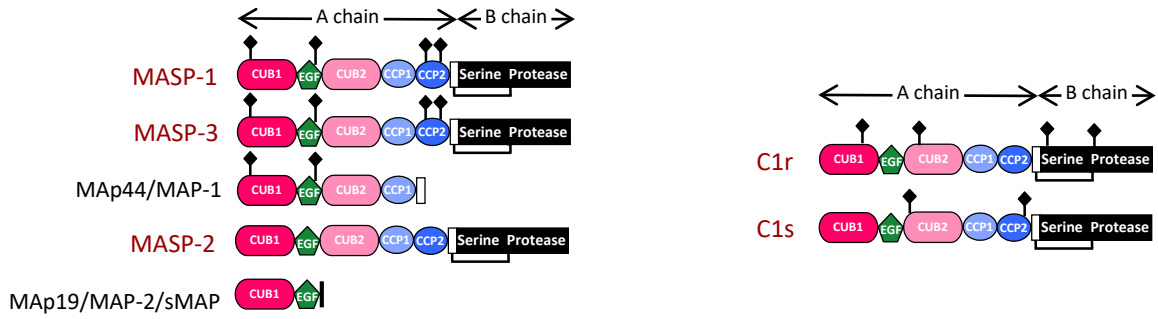
46. Gaboriaud C, Ling WL, Thielens NM, Bally I, Rossi V. Deciphering the fine details of c1 assembly and activation mechanisms: “mission impossible”? *Front Immunol* (2014) **5**:565. doi:10.3389/fimmu.2014.00565
47. Tschopp J, Villiger W, Fuchs H, Kilchherr E, Engel J. Assembly of subcomponents C1r and C1s of first component of complement: electron microscopic and ultracentrifugal studies. *Proc Natl Acad Sci USA* (1980) **77**:7014–7018. doi:10.1073/pnas.77.12.7014
48. Goettig P, Brandstetter H, Magdolen V. Surface loops of trypsin-like serine proteases as determinants of function. *Biochimie* (2019) **166**:52–76. doi:10.1016/j.biochi.2019.09.004
49. Chakraborty P, Acquasaliente L, Pelc LA, Di Cera E. Interplay between conformational selection and zymogen activation. *Sci Rep* (2018) **8**:4080. doi:10.1038/s41598-018-21728-9
50. Matthews BW, Sigler PB, Henderson R, Blow DM. Three-dimensional structure of tosyl-alpha-chymotrypsin. *Nature* (1967) **214**:652–656. doi:10.1038/214652a0
51. Hartley BS. Homologies in serine proteinases. *Philos Trans R Soc Lond, B, Biol Sci* (1970) **257**:77–87. doi:10.1098/rstb.1970.0010
52. Drentin N, Conroy P, Gunzburg MJ, Pike RN, Wijeyewickrema LC. Investigation of the mechanism of interaction between Mannose-binding lectin-associated serine protease-2 and complement C4. *Mol Immunol* (2015) **67**:287–293. doi:10.1016/j.molimm.2015.06.011
53. Perry AJ, Wijeyewickrema LC, Wilmann PG, Gunzburg MJ, D’Andrea L, Irving JA, Pang SS, Duncan RC, Wilce JA, Whisstock JC, et al. A molecular switch governs the interaction between the human complement protease C1s and its substrate, complement C4. *J Biol Chem* (2013) **288**:15821–15829. doi:10.1074/jbc.M113.464545
54. Perona JJ, Craik CS. Structural basis of substrate specificity in the serine proteases. *Protein Sci* (1995) **4**:337–360. doi:10.1002/pro.5560040301
55. Gaboriaud C, Gupta RK, Martin L, Lacroix M, Serre L, Teillet F, Arlaud GJ, Rossi V, Thielens NM. The serine protease domain of MASP-3: enzymatic properties and crystal structure in complex with ecotin. *PLoS ONE* (2013) **8**:e67962. doi:10.1371/journal.pone.0067962
56. Pang SS, Wijeyewickrema LC, Hor L, Tan S, Lameignere E, Conway EM, Blom AM, Mohlin FC, Liu X, Payne RJ, et al. The Structural Basis for Complement Inhibition by Gigastasin, a Protease Inhibitor from the Giant Amazon Leech. *J Immunol* (2017) **199**:3883–3891. doi:10.4049/jimmunol.1700158
57. Nagy ZA, Szakács D, Boros E, Héja D, Vígh E, Sándor N, Józsi M, Oroszlán G, Dobó J, Gál P, et al. Ecotin, a microbial inhibitor of serine proteases, blocks multiple complement dependent and independent microbicidal activities of human serum. *PLoS Pathog* (2019) **15**:e1008232. doi:10.1371/journal.ppat.1008232
58. Dobó J, Szakács D, Oroszlán G, Kortvely E, Kiss B, Boros E, Szász R, Závodszy P, Gál P, Pál G. MASP-3 is the exclusive pro-factor D activator in resting blood: the lectin and the alternative complement pathways are fundamentally linked. *Sci Rep* (2016) **6**:31877. doi:10.1038/srep31877
59. Héja D, Harmat V, Fodor K, Wilmanns M, Dobó J, Kékesi KA, Závodszy P, Gál P, Pál G. Monospecific inhibitors show that both mannan-binding lectin-associated serine protease-1 (MASP-1) and -2 Are essential for lectin pathway activation and reveal structural plasticity of MASP-2. *J Biol Chem* (2012) **287**:20290–20300. doi:10.1074/jbc.M112.354332

60. Héja D, Kocsis A, Dobó J, Szilágyi K, Szász R, Závodszy P, Pál G, Gál P. Revised mechanism of complement lectin-pathway activation revealing the role of serine protease MASP-1 as the exclusive activator of MASP-2. *Proc Natl Acad Sci USA* (2012) **109**:10498–10503. doi:10.1073/pnas.1202588109
61. Dobó J, Pál G, Cervenak L, Gál P. The emerging roles of mannose-binding lectin-associated serine proteases (MASPs) in the lectin pathway of complement and beyond. *Immunol Rev* (2016) **274**:98–111. doi:10.1111/imr.12460
62. Degn SE, Kjaer TR, Kidmose RT, Jensen L, Hansen AG, Tekin M, Jensenius JC, Andersen GR, Thiel S. Complement activation by ligand-driven juxtaposition of discrete pattern recognition complexes. *Proc Natl Acad Sci USA* (2014) **111**:13445–13450. doi:10.1073/pnas.1406849111
63. Mortensen SA, Sander B, Jensen RK, Pedersen JS, Golas MM, Jensenius JC, Hansen AG, Thiel S, Andersen GR. Structure and activation of C1, the complex initiating the classical pathway of the complement cascade. *Proc Natl Acad Sci USA* (2017) **114**:986–991. doi:10.1073/pnas.1616998114
64. Paréj K, Dobó J, Závodszy P, Gál P. The control of the complement lectin pathway activation revisited: both C1-inhibitor and antithrombin are likely physiological inhibitors, while α 2-macroglobulin is not. *Mol Immunol* (2013) **54**:415–422. doi:10.1016/j.molimm.2013.01.009
65. Ambrus G, Gál P, Kojima M, Szilágyi K, Balczer J, Antal J, Gráf L, Laich A, Moffatt BE, Schwaeble W, et al. Natural substrates and inhibitors of mannan-binding lectin-associated serine protease-1 and -2: a study on recombinant catalytic fragments. *J Immunol* (2003) **170**:1374–1382. doi:10.4049/jimmunol.170.3.1374
66. Garred P, Genster N, Pilely K, Bayarri-Olmos R, Rosbjerg A, Ma YJ, Skjoedt M-O. A journey through the lectin pathway of complement-MBL and beyond. *Immunol Rev* (2016) **274**:74–97. doi:10.1111/imr.12468
67. Yongqing T, Wilmann PG, Reeve SB, Coetzer TH, Smith AI, Whisstock JC, Pike RN, Wijeyewickrema LC. The x-ray crystal structure of mannose-binding lectin-associated serine proteinase-3 reveals the structural basis for enzyme inactivity associated with the Carnevale, Mingarelli, Malpuech, and Michels (3MC) syndrome. *J Biol Chem* (2013) **288**:22399–22407. doi:10.1074/jbc.M113.483875
68. Atik T, Koparir A, Bademci G, Foster J, Altunoglu U, Mutlu GY, Bowdin S, Elcioglu N, Tayfun GA, Atik SS, et al. Novel MASP1 mutations are associated with an expanded phenotype in 3MC1 syndrome. *Orphanet J Rare Dis* (2015) **10**:128. doi:10.1186/s13023-015-0345-3
69. Graul-Neumann LM, Mensah MA, Klopocki E, Uebe S, Ekici AB, Thiel CT, Reis A, Zweier C. Biallelic intragenic deletion in MASP1 in an adult female with 3MC syndrome. *Eur J Med Genet* (2018) **61**:363–368. doi:10.1016/j.ejmg.2018.01.016
70. Munye MM, Diaz-Font A, Ocaka L, Henriksen ML, Lees M, Brady A, Jenkins D, Morton J, Hansen SW, Bacchelli C, et al. COLEC10 is mutated in 3MC patients and regulates early craniofacial development. *PLoS Genet* (2017) **13**:e1006679. doi:10.1371/journal.pgen.1006679
71. Kapferer-Seebacher I, Pepin M, Werner R, Aitman TJ, Nordgren A, Stoiber H, Thielens N, Gaboriaud C, Amberger A, Schossig A, et al. Periodontal Ehlers-Danlos Syndrome Is Caused by Mutations in C1R and C1S, which Encode Subcomponents C1r and C1s of Complement. *Am J Hum Genet* (2016) **99**:1005–1014. doi:10.1016/j.ajhg.2016.08.019

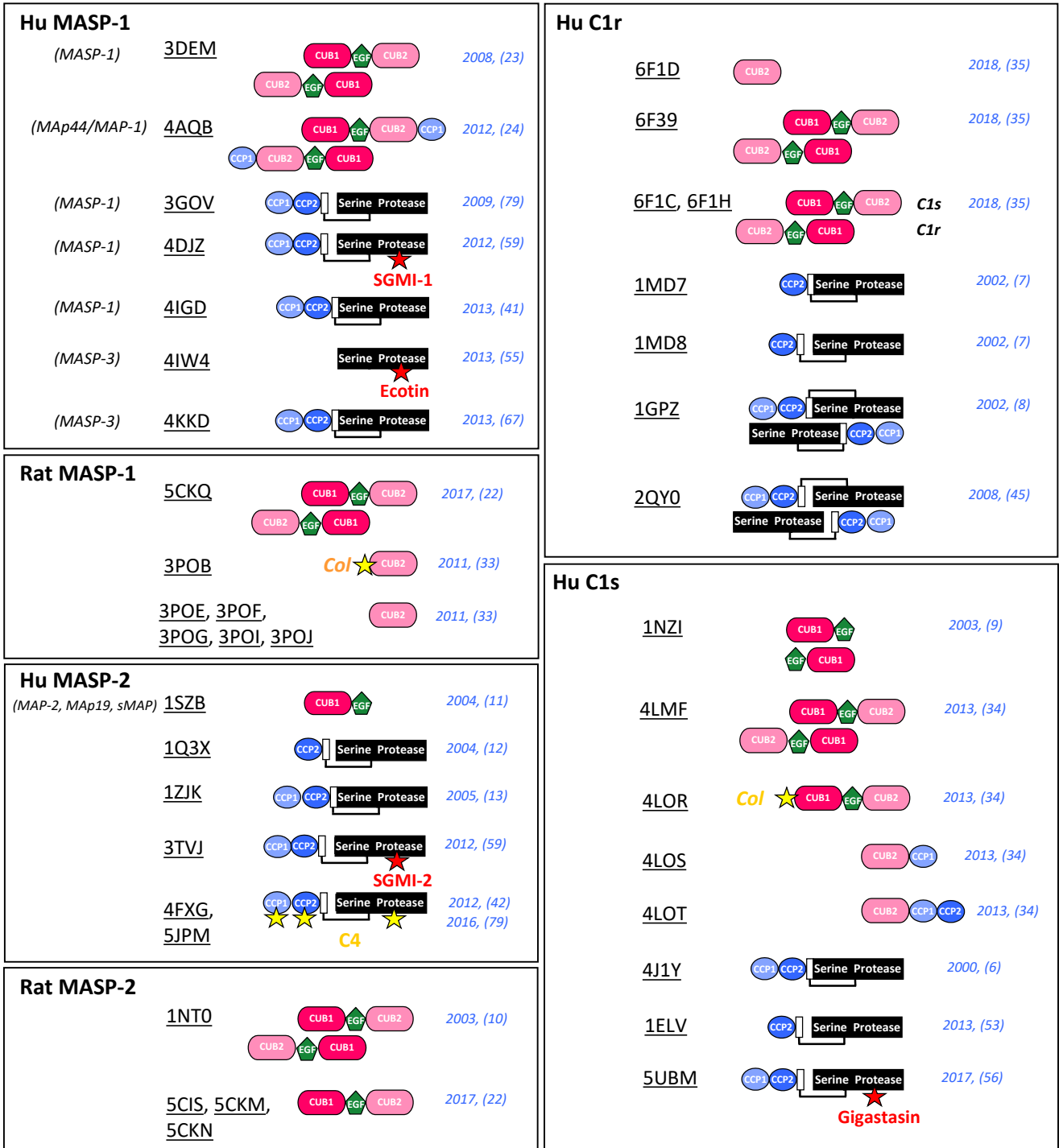
72. Gröbner R, Kapferer-Seebacher I, Amberger A, Redolfi R, Dalonneau F, Björck E, Milnes D, Bally I, Rossi V, Thielens N, et al. C1R Mutations Trigger Constitutive Complement 1 Activation in Periodontal Ehlers-Danlos Syndrome. *Front Immunol* (2019) **10**:2537. doi:10.3389/fimmu.2019.02537
73. Bally I, Dalonneau F, Chouquet A, Gröbner R, Amberger A, Kapferer-Seebacher I, Stoiber H, Zschocke J, Thielens NM, Rossi V, et al. Two Different Missense C1S Mutations, Associated to Periodontal Ehlers-Danlos Syndrome, Lead to Identical Molecular Outcomes. *Front Immunol* (2019) **10**:2962. doi:10.3389/fimmu.2019.02962
74. Hajishengallis G, Abe T, Maekawa T, Hajishengallis E, Lambris JD. Role of complement in host-microbe homeostasis of the periodontium. *Semin Immunol* (2013) **25**:65–72. doi:10.1016/j.smim.2013.04.004
75. Krem MM, Di Cera E. Molecular markers of serine protease evolution. *EMBO J* (2001) **20**:3036–3045. doi:10.1093/emboj/20.12.3036
76. Dobó J, Schroeder V, Jenny L, Cervenak L, Závodszky P, Gál P. Multiple roles of complement MASP-1 at the interface of innate immune response and coagulation. *Mol Immunol* (2014) **61**:69–78. doi:10.1016/j.molimm.2014.05.013
77. Gál P, Dobó J, Závodszky P, Sim RBM. Early complement proteases: C1r, C1s and MASPs. A structural insight into activation and functions. *Mol Immunol* (2009) **46**:2745–2752. doi:10.1016/j.molimm.2009.04.026
78. Gohara DW, Di Cera E. Allostery in trypsin-like proteases suggests new therapeutic strategies. *Trends Biotechnol* (2011) **29**:577–585. doi:10.1016/j.tibtech.2011.06.001
79. Croll TI, Andersen GR. Re-evaluation of low-resolution crystal structures via interactive molecular-dynamics flexible fitting (iMDFF): a case study in complement C4. *Acta Crystallogr D Struct Biol* (2016) **72**:1006–1016. doi:10.1107/S2059798316012201
80. Dobó J, Harmat V, Beinrohr L, Sebestyén E, Závodszky P, Gál P. MASP-1, a promiscuous complement protease: structure of its catalytic region reveals the basis of its broad specificity. *J Immunol* (2009) **183**:1207–1214. doi:10.4049/jimmunol.0901141

Fig.1

A



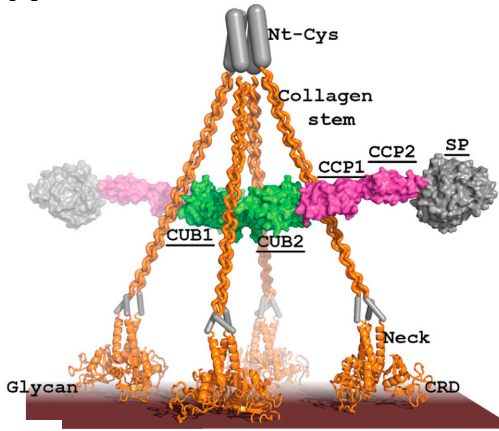
B



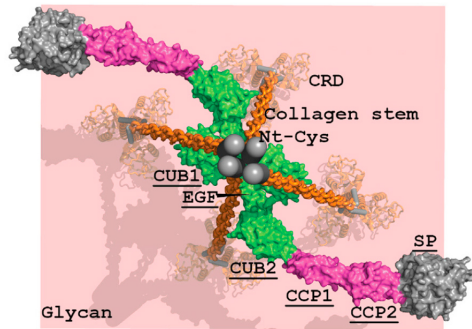
MASP-1, Hu PO:48740-1, Rat PO:Q8CHN8; MASP-3, Hu PO: 48740-2; MAP44/MAP-1, Hu PO:48740-3; C1r, Hu PO: P00736
 MASP-2, Hu PO:O00187-1, Rat PO:Q9JJS8; MAP19/MAP-2/sMAP, Hu PO:O10187-2; C1s, Hu PO: P09871

Fig.2

A



B



C

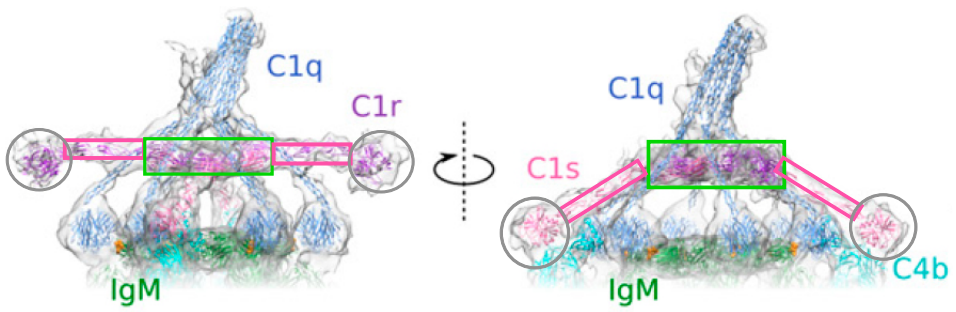


Fig.3

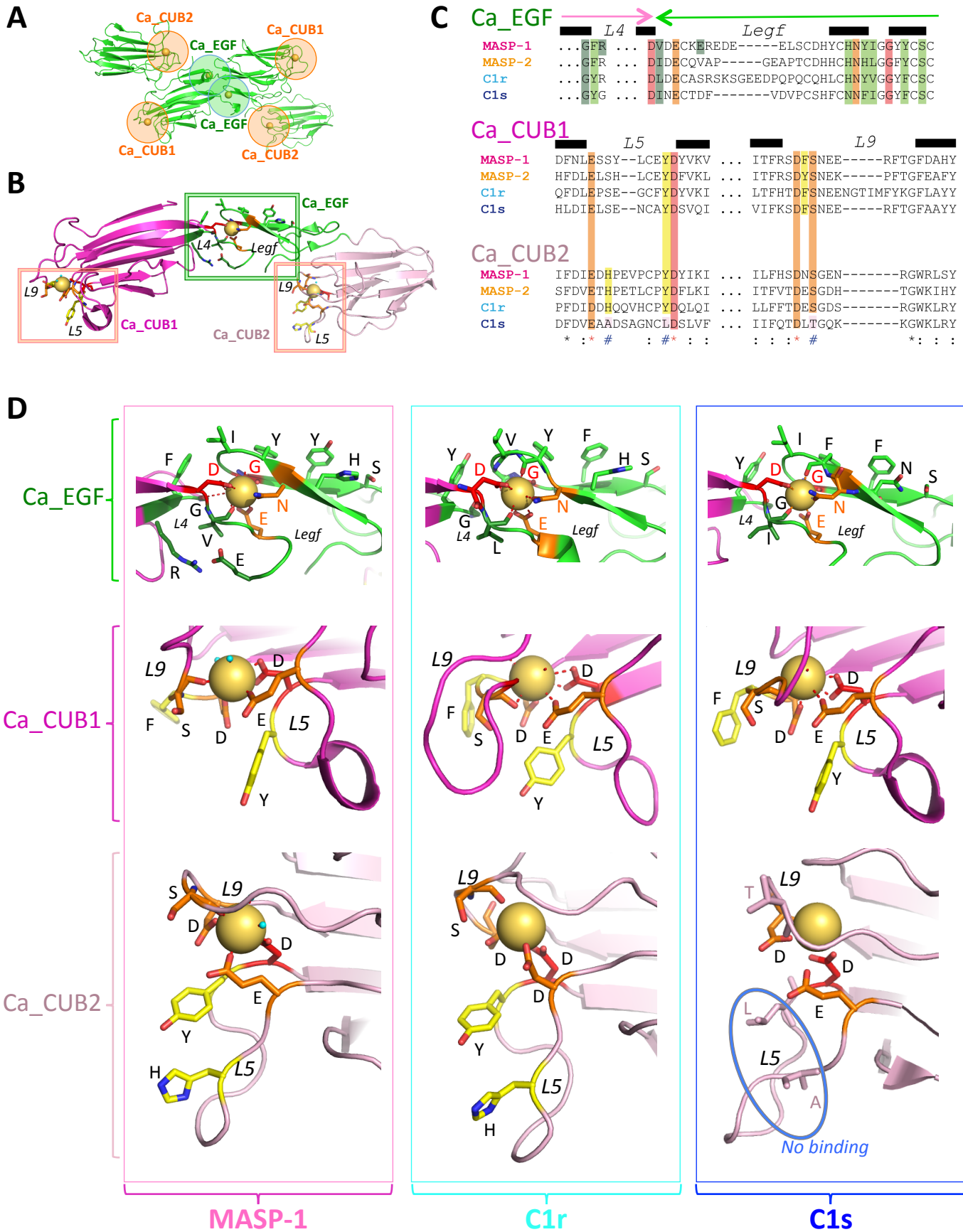


Fig.4

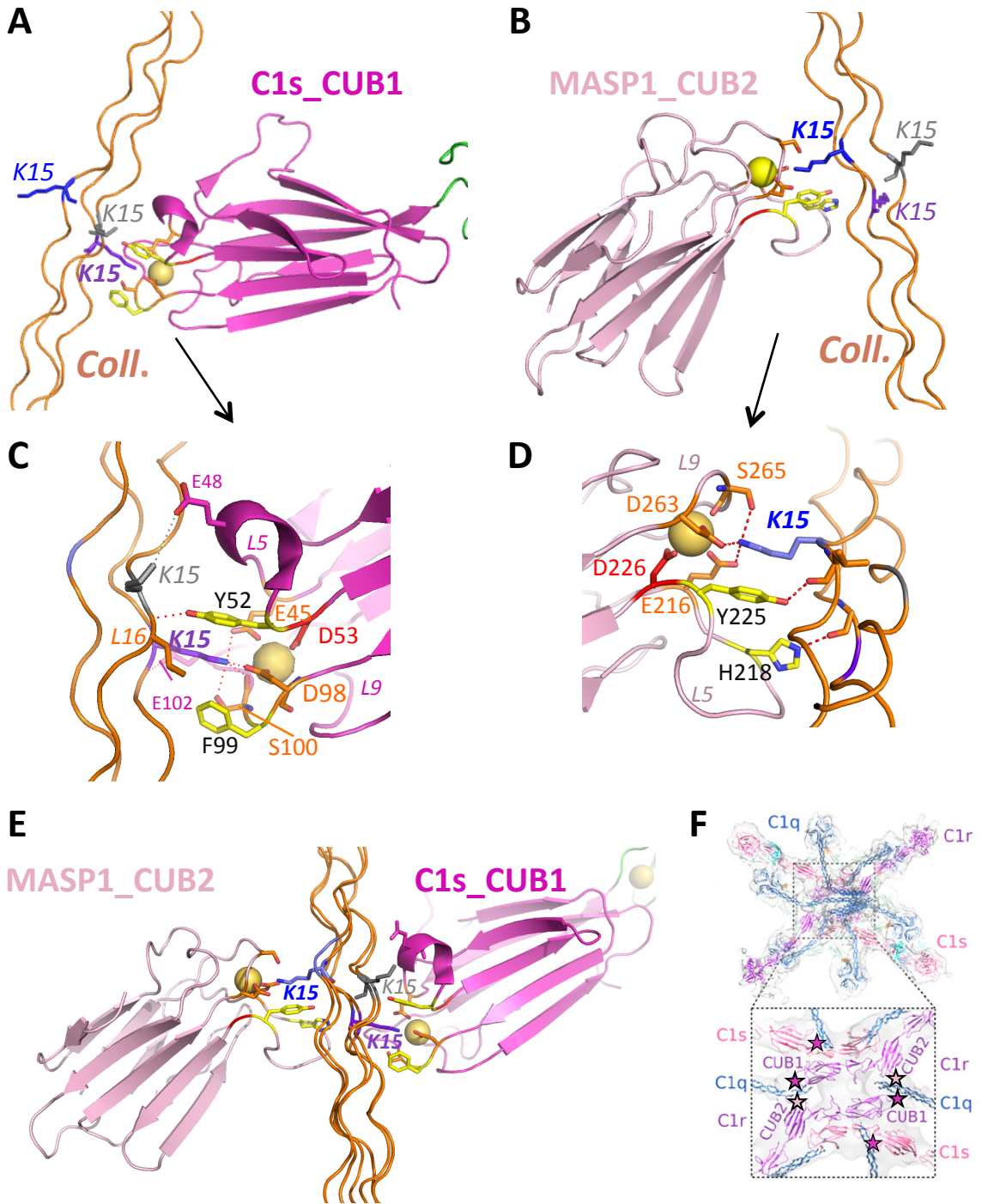
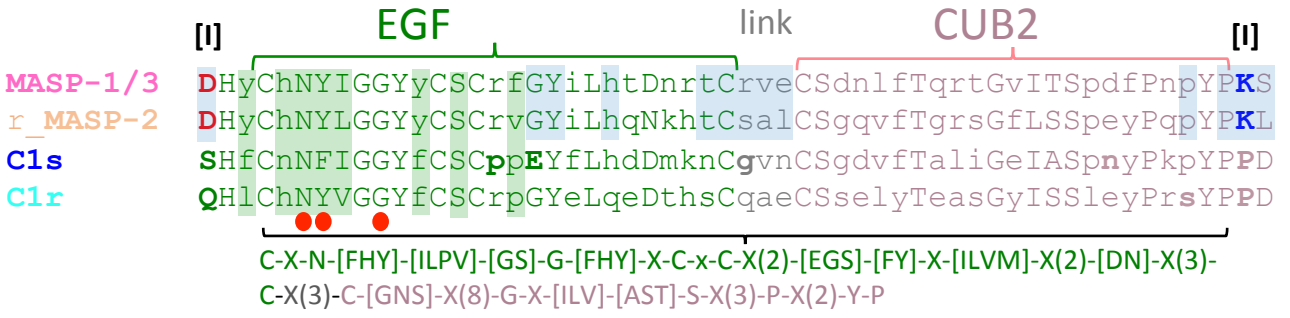
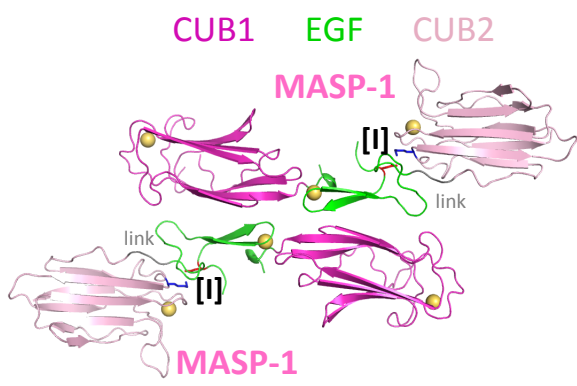


Fig.5

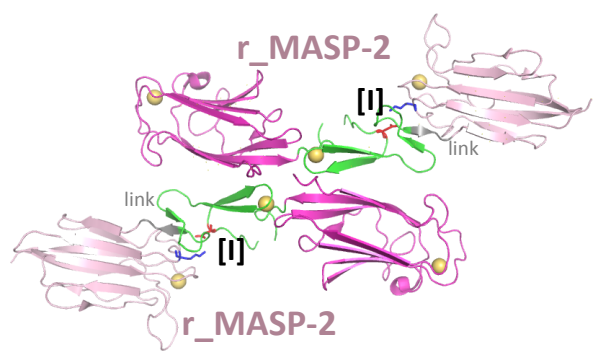
A



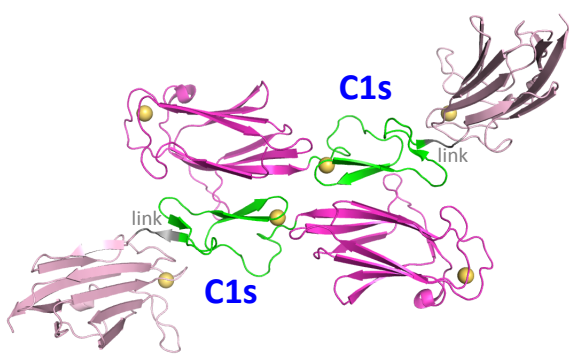
B



C



D



E

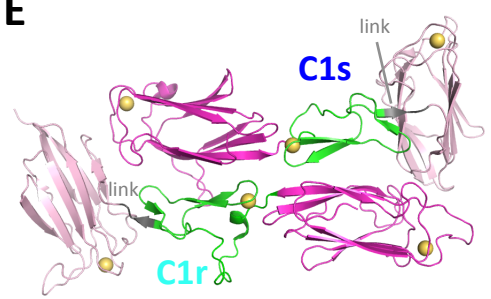


Fig.6

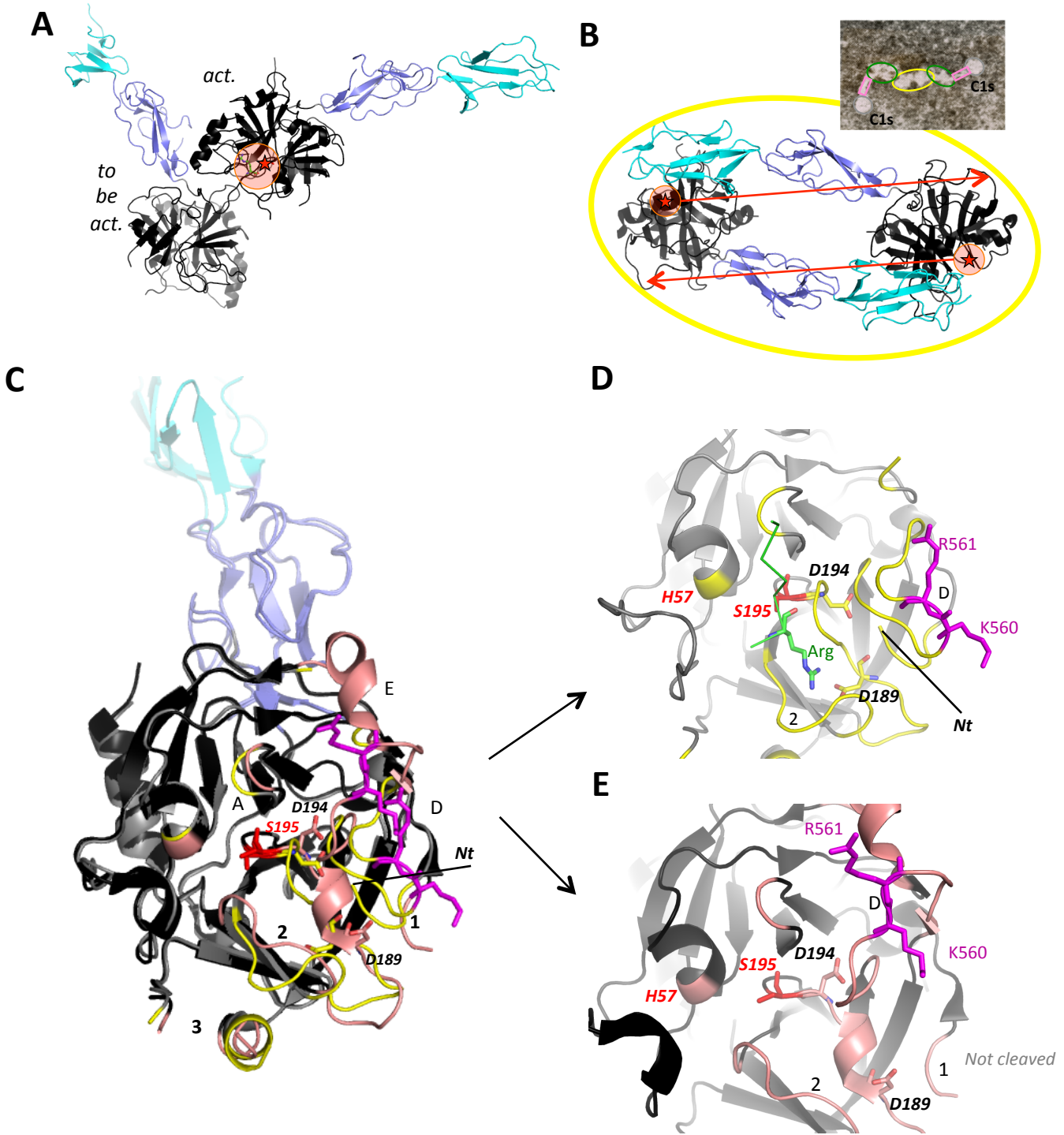


Fig.7

