Review

Theodoros Goulas, Irene Garcia-Ferrer[a], Aniebrys Marrero[b], Laura Marino-Puertas, Stephane Duquerroy and F. Xavier Gomis-Rüth*

Structural and functional insight into pan-endopeptidase inhibition by α_2-macroglobulins

DOI 10.1515/hsz-2016-0329
Received November 22, 2016; accepted February 18, 2017; previously published online March 1, 2017

Abstract: Peptidases must be exquisitely regulated to prevent erroneous cleavage and one control is provided by protein inhibitors. These are usually specific for particular peptidases or families and sterically block the active-site cleft of target enzymes using lock-and-key mechanisms. In contrast, members of the +1400-residue multi-domain α_2-macroglobulin inhibitor family (α_2Ms) are directed against a broad spectrum of endopeptidases of disparate specificities and catalytic types, and they inhibit their targets without disturbing their active sites. This is achieved by irreversible trap mechanisms resulting from large conformational rearrangement upon cleavage in a promiscuous bait region through the prey endopeptidase. After decades of research, high-resolution structural details of these mechanisms have begun to emerge for tetrameric and monomeric α_2Ms, which use 'Venus-flytrap' and 'snap-trap' mechanisms, respectively. In the former, represented by archetypal human α_2M, inhibition is exerted through physical entrapment in a large cage, in which preys are still active against small substrates and inhibitors that can enter the cage through several apertures. In the latter, represented by a bacterial α_2M from *Escherichia coli*, covalent linkage and steric hindrance of the prey inhibit activity, but only against very large substrates.

Keywords: bait region cleavage; conformational rearrangement; irreversible inhibition; multi-domain protein; protein inhibitor; regulation of proteolytic activity.

[a]Present address: EMBL Grenoble, 71 Avenue des Martyrs, CS 90181, F-38042 Grenoble Cedex 9, France
[b]Present address: Department of Chemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
***Corresponding author: F. Xavier Gomis-Rüth,** Proteolysis Laboratory, Structural Biology Unit, 'María-de-Maeztu' Unit of Excellence, Molecular Biology Institute of Barcelona (CSIC), Barcelona Science Park, c/Baldiri Reixac, 15-21, E-08028 Barcelona, Spain, e-mail: xgrcri@ibmb.csic.es
Theodoros Goulas, Irene Garcia-Ferrer, Aniebrys Marrero and Laura Marino-Puertas: Proteolysis Laboratory, Structural Biology Unit, 'María-de-Maeztu' Unit of Excellence, Molecular Biology Institute of Barcelona (CSIC), Barcelona Science Park, c/Baldiri Reixac, 15-21, E-08028 Barcelona, Spain
Stephane Duquerroy: Unité de Virologie Structurale, Département de Virologie and CNRS Unité de Recherche Associée (URA) 3015, Institut Pasteur, 25 Rue du Dr Roux, F-75724 Paris Cedex 15, France; and Departement de Biologie, Faculté des Sciences d'Orsay, Université Paris-Sud, F-91405 Orsay, France

Introduction: the α_2-macroglobulin family of inhibitors

Proteolysis is carried out by peptidases and is a universal mechanism of post-translational modification of proteins and peptides. It generates active forms from latent precursors or inactivates proteins that are obsolete or defective (Neurath and Walsh, 1976). In addition, proteolysis causes widespread degradation of proteins and peptides during digestion and tissue remodeling. Proteolysis is also observed in microbial infections, where it either turns off defense proteins from the attacked organism and provides nutrients for the infecting microbe or provides a first line of innate host defense response (Armstrong, 2006; Dubin et al., 2013). Proteolysis entails cleavage of peptide bonds, which is generally irreversible. Therefore, it must be exquisitely regulated, both temporally and spatially, to prevent aberrant reactions. Failure in this regulation is observed in all major pathologies, ranging from inflammation, tissue destruction and neurological disorders to cardiovascular diseases, diabetes and cancer. Peptidases are kept in check by protein inhibitors, among other regulatory mechanisms (Travis and Salvesen, 1983). While prokaryotes in general contain only a few inhibitors, if any (Kantyka et al., 2010), animals, plants and viruses produce many of them, which are grouped in +75 families according to the MEROPS database (http://merops.sanger.ac.uk; Rawlings et al., 2016). One of the families comprises the

α_2-macroglobulins (α_2Ms; MEROPS family I39), which are large multi-domain proteins that act as broad-spectrum endopeptidase inhibitors (Barrett and Starkey, 1973; Rehman et al., 2013). They are distantly related to and possess similar overall architecture and functional relationship to innate-immunity proteins from the complement system and related thioester-containing proteins, but these are not peptidase inhibitors (Sottrup-Jensen et al., 1985; Armstrong and Quigley, 1999; Vogel et al., 2014).

The founding and best-studied member of the α_2Ms is human α_2M [hα_2M; (Barrett and Starkey, 1973; Starkey and Barrett, 1973; Barrett, 1981)], which was first isolated by Cohn and colleagues in 1946 (Cohn et al., 1946) and named by Schultze and co-workers in 1955 (Schultze et al., 1955). Other human α_2M members studied are pregnancy zone protein [PZP; (Sand et al., 1985)] and protein α_2M-like 1 [α2ML1; (Galliano et al., 2006)]. Moreover, α_2Ms have been reported in other animals including mammals, reptiles, amphibians, fishes and invertebrates (Starkey and Barrett, 1982; Sottrup-Jensen, 1989; Stöcker et al., 1991; Rubenstein et al., 1993; Bender and Bayne, 1996; Armstrong and Quigley, 1999; Armstrong, 2006; Chaikeeratisak et al., 2012). In addition, ovostatins (*alias* ovomacroglobulins) are found in avian and reptilian egg whites (Nagase and Harris Jr., 1983; Woessner Jr. and Nagase, 2000).

Outside metazoans, genes similar to α_2Ms are generally absent from archaea, protozoa, fungi and plants. Strikingly, they are found in colonizing Gram-negative proteobacteria but not in free-living or environmental microbes. Their presence in pathogenic species and strains such as *Salmonella typhimurium*, *Rickettsia prowazekii*, *Yersinia pestis*, *Pseudomonas aeruginosa* and enteropathogenic, enterotoxigenic and enterohemorrhagic *Escherichia coli* make them attractive targets for the development of antimicrobials (Budd et al., 2004; Doan and Gettins, 2008; Neves et al., 2012; Robert-Genthon et al., 2013; Wong and Dessen, 2014; Garcia-Ferrer et al., 2015). Noteworthy, phylogenetic studies revealed that the gene distribution in bacteria is patchy and violates the vertical descent model, which suggests that bacterial α_2Ms (bα_2Ms) may be xenologs acquired from metazoans by horizontal gene transfer facilitated by long-standing intimate contact between colonizing bacteria and their warm-blooded animal hosts (Budd et al., 2004; Doan and Gettins, 2008).

As to the oligomerization state of α_2Ms, some inhibitors have been reported to be dimeric, such as human PZP (Sand et al., 1985) and forms from several crustacea, cephalopoda and fishes (Starkey and Barrett, 1982; Sottrup-Jensen, 1989; Armstrong, 2006). Others are tetrameric and monomeric, as discussed below.

Biochemical mechanism of tetrameric human α_2M inhibitor

Authentic hα_2M is synthesized in the liver as a 1,474-residue protein provided with a 23-residue signal peptide for secretion, and it is polyglycosylated (Armstrong and Quigley, 1999). It is found in the circulation at high concentration (~3% of the total plasma proteins in humans; Ganrot and Schersten, 1967) as a ~720-kDa homotetramer, which is actually a dimer of disulfide-linked dimers (Barrett et al., 1979). Human α_2M targets a large variety of endopeptidases of the four major classes regardless of their architecture and catalytic type (Barrett, 1981). Seven decades of painstaking biochemistry by dozens of groups worldwide revealed that inhibition by hα_2M does not occur through an exposed flexible segment targeting and blocking the active site of endopeptidases, as found, for example, in the mostly specific 'standard-mechanism' inhibitors (Laskowski Jr. and Kato, 1980; Enghild et al., 1989a; Bode and Huber, 1992; Laskowski Jr. and Qasim, 2000; Arolas et al., 2011). Instead, it operates through a unique irreversible 'Venus-flytrap' (Conrad, 1998; Marrero et al., 2012) or 'trap-hypothesis' mechanism (Barrett and Starkey, 1973; see Figure 1A and C), which once sprung encages peptidases, but does not block them so they remain active against small substrates or inhibitors. However, action on medium-sized and large substrates and inhibitors is sterically hindered (Barrett and Starkey, 1973). In general, small peptidases the size of HIV proteinase or trypsin react rapidly and form 2:1 complexes with tetrameric hα_2M, while large peptidases the size of plasmin or very specific peptidases react slowly and form 1:1 complexes (Barrett and Starkey, 1973; Travis and Salvesen, 1983; Sottrup-Jensen, 1989).

In its open, unreacted 'native' conformation, the hα_2M tetramer can accommodate prey peptidases inside it (Figure 1A, left; Marrero et al., 2012). There, peptidases cleave the inhibitor within a 'bait region', which is located nearly in the middle of the polypeptide chain in a flexible and accessible 'bait region domain' (BRD; for domain nomenclature of hα_2M, see Figure 2A). The bait region differs among mammalian α_2Ms in sequence and length (~30–50 residues; Sottrup-Jensen et al., 1989) and contains many potential recognition sites, so endopeptidases that do not cleave either have rare specificity or are too large to be internalized by the tetramer (Ikai et al., 1999).

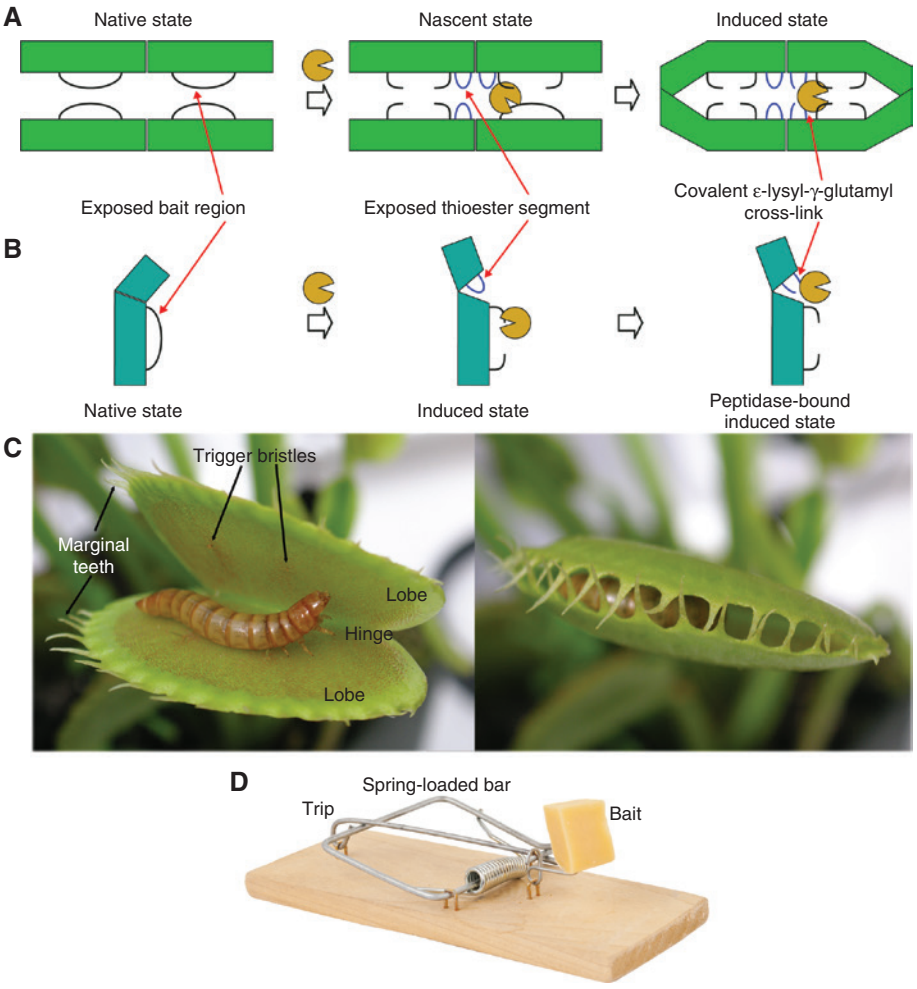

Figure 1: Mechanisms of α_2Ms.
(A) Schematic mechanism of endopeptidase inhibition through tetrameric hα_2M. The native state is open and the bait region is exposed and accessible for a prey peptidase (yellow Pac-Man). Once the latter cleaves the bait region, the inhibitor enters a nascent state, in which the thioester-segment – which was hidden in the native state – is exposed and becomes targetable by lysines on the surface of the prey peptidase. This gives rise to a covalent cross-link. Overall, conformational rearrangement leads to the closed induced state of the tetramer. (B) Schematic mechanism of endopeptidase inhibition through monomeric ECAM. In contrast to panel (A), inhibition only occurs after cleavage in the bait region if the prey peptidase is bound through a covalent linkage mediated by the thioester bond. (C) The Venus flytrap *Dionaea muscipula* closes when the prey touches trigger bristles on the inner surface of either lobe, which act as bait (left). Special motor cells in the hinge between the lobes are filled with liquid in the triggered state. When the trigger is fired, the liquid is expelled from the cells, which causes the trap to spring shut in a fifth of a second, so that the lobes close over the prey. The cilia or marginal teeth surrounding the lobes point slightly outwards and interdigitate (right). Inside the closed trap, the prey can still be reached through openings between the marginal teeth. Photos of a meal worm in a Venus flytrap, courtesy of Beatrice Murch (http://www.beatricemurchphotography.com). (D) Baited and set snap trap. The spring-loaded bar is held down by a trip. Bait manipulation leads to trip displacement, and the trap is set off through a rapid down swing of the spring-loaded bar. Image downloaded from http://disruptiveviews.com/fraudsters-caught-by-neural-science with free permission to share.

In addition, the fact that the bait region is imbedded within a continuous polypeptide chain means that no exopeptidases are inhibited, as these require free N- or C-termini for catalysis. Cleavage leads to a 'nascent state' of α_2M, in which a buried, very reactive β-cysteinyl-γ-glutamyl 'thioester bond' within a 'thioester domain' (TED) becomes exposed (Figure 1A, center; Sottrup-Jensen et al., 1981b, 1989; Sottrup-Jensen, 1989). The thioester bond results from post-translational covalent linkage between the side chains of a cysteine and a glutamate/glutamine three positions downstream in the sequence (Iijima et al., 1984; Grøn et al., 1996). The bond is readily cleaved by amines from surface lysines of the prey endopeptidase (Armstrong and Quigley, 1999), so it must be shielded in native

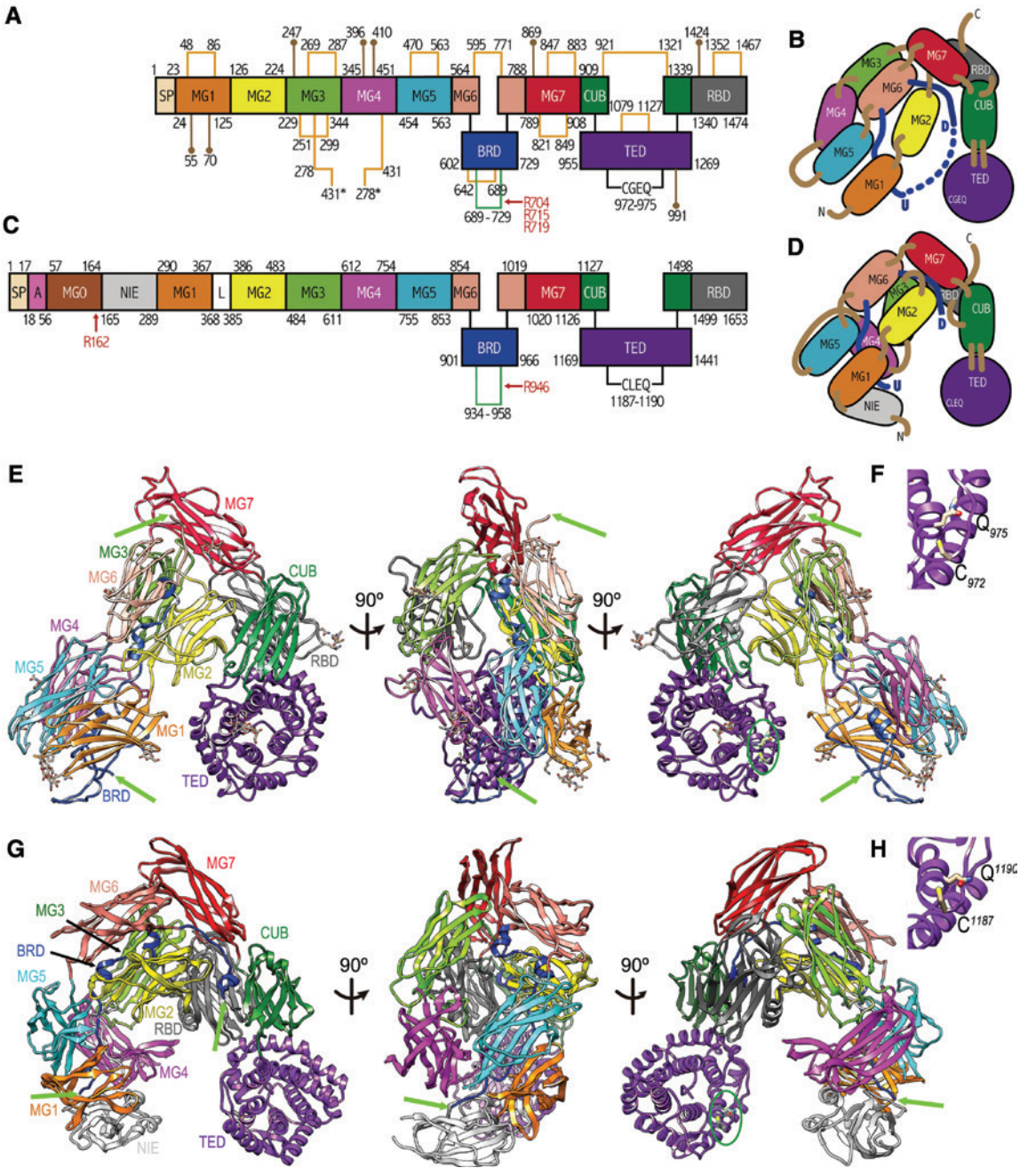

Figure 2: Protomers of α_2Ms.
(A) Domain structure of 1,474-residue hα_2M (UniProt database [UP] access code P01023). Each domain is shown in one color, together with the respective spanning residues (SP, signal peptide; MG1-MG7, macroglobulin domains 1–7; BRD, bait-region domain; CUB domain; TED, thioester domain; and RBD, receptor-binding domain). Disulfide bonds are shown in orange and the connected cysteines are labeled. Two such bonds symmetrically connect two protomers within a dimer through respective residues C_{431} and C_{278}. N-glycosylated residues are pinpointed by brown lollipops and labeled. The bait region within BRD is shown in green, the thioester segment within TED in black. Within the bait region, trypsin cleavage sites are shown in red (see Sottrup-Jensen et al., 1981a). (B) Spatial domain arrangement of MA-induced hα_2M in the front view chosen for reference. Domains are connected with brown coils. BRD is represented by a dark-blue coil, with the flexible but intact bait region in a discontinuous trace. The segments upstream (U) and downstream (D) of the bait region are labeled. (C) Domain structure of 1653-residue ECAM (UP P76578). In addition to the domains found in hα_2M, ECAM possesses a flexible anchor to the membrane (A), an extra MG domain (MG0), an N-terminal domain of induced ECAM (NIE) and a linker (L). Red arrows pinpoint the primary trypsin cleavage sites for induction (R^{946}) and shedding/dimerization (R^{162}). (D) Scheme of iECAM in the reference front view as in (B). Here, the bait region is cleaved. (E) Richardson-plot of the MA-induced hα_2M protomer in front (left panel), lateral (middle panel), and back views (right panel). Green arrows point to the visible ends of the flexible but intact bait region in the experimental structure. In the right panel, a green ellipse pinpoints the thioester motif within TED, which is blown up in (F). (G, H) Same as (E) and (F) for the ECAM monomer.

forms of thioester proteins to prevent precocious opening (Blandin and Levashina, 2004). Thioester cleavage results in a free cysteine and an ε-lysyl-γ-glutamyl cross-link between the side chains of the thioester glutamine and the prey lysine (Salvesen et al., 1981; Sottrup-Jensen, 1994). However, some peptidases are efficiently inhibited by tetrameric α_2Ms but are not covalently bound due to the absence of surface lysines, such as human neutrophil elastase (Enghild et al., 1989a), or the rapid solvent-mediated decay of the thioester bond once exposed (Wang et al., 1983). Consistently, ovostatin from chicken egg is an α_2M-like inhibitor of similar efficiency to hα_2M but it lacks the thioester bond (Nagase and Harris Jr., 1983; Nagase et al., 1983).

Cleavage in the bait region also causes large structural rearrangement of the tetrameric particle. This gives rise to a closed, reacted, activated or 'induced' state, which results in peptidase entrapment in a similar manner to insects that are trapped by the carnivorous Venus flytrap (Figure 1A and C; Barrett and Starkey, 1973). The native and induced states can be distinguished through different sedimentation coefficients (Björk and Fish, 1982) and mobilities on native gel electrophoresis, as they are 'slow' and 'fast' species, respectively (Barrett et al., 1979). Inside the closed cage, prey still cleave small substrates (Travis and Salvesen, 1983) and interact with inhibitors up to ~10–20 kDa of mass (Bieth et al., 1981). A similar conformational rearrangement leading to induction can be obtained by reaction of the thioester bond with small primary amines such as methylamine (MA). This reaction is slower than proteolytic induction and yields a free cysteine thiol group plus an N-substituted glutamine. In this form of tetrameric hα_2M, the bait regions are intact but the molecule no longer reacts with and inhibits endopeptidases because the trap is closed (Sottrup-Jensen et al., 1980; Travis and Salvesen, 1983; Armstrong and Quigley, 1999; Qazi et al., 1999). Induction of hα_2M further exposes the C-terminal receptor-binding domain (RBD) of each protomer on the surface of the particle (Delain et al., 1988). RBDs are bound by specific cell-surface receptors such as the low-density lipoprotein receptor-related protein found on fibroblasts, macrophages, syncytiotrophoblasts, hepatocytes and other cell types (Andersen et al., 2000). This triggers clearance of the inhibitor and its prey from the circulation – within minutes of its formation – via receptor-mediated endocytosis and lysosomal degradation (Imber and Pizzo, 1981; Strickland et al., 1990). In native hα_2M, the RBDs are hidden and cannot interact with the receptors (Debanne et al., 1976; Strickland et al., 1990).

Overall, other tetrameric α_2Ms likely to operate similarly to hα_2M have been described from mammals (bovine, rat, several monkeys, hedgehog, squirrel, tiger, elephant, camel and wallaby), birds (goose, chicken and duck), amphibians (several frogs), reptiles (several snakes and alligator) and invertebrates (snail) (Starkey and Barrett, 1982; Sottrup-Jensen, 1989; Bender and Bayne, 1996; Armstrong and Quigley, 1999). Also egg ovostatins likely proceed in a similar fashion (Nagase and Harris Jr., 1983; Nagase et al., 1983).

Structural aspects of tetrameric α_2M inhibitors

Tetrameric α-macroglobulins have been the object of structural studies for decades, and several crystallization reports have been published (Brown et al., 1954; Andersen et al., 1991, 1994; Goulas et al., 2014). However, due to the intrinsic flexibility and heterogeneity of authentic purified proteins, only low-resolution electron microscopy reconstructions were available for a long time, which in general confirmed the major structural differences between native and induced tetramers predicted by biophysical studies (Delain et al., 1988, 1992, 1994; Stoops et al., 1991; Boisset et al., 1993, 1996; Kolodziej et al., 1996, 2002; Qazi et al., 1999). In addition, a Fourier map of MA-induced hα_2M from crystal diffraction to 10 Å was reported in 1995 (Andersen et al., 1995). High- or medium-resolution structures could only be obtained for the isolated α_2M RBDs from human (Protein Data Bank [PDB] access code 1BV8; Huang et al., 2000) and bovine (PDB 1AYO; Jenner et al., 1998), for the RBD from rat α_1M [PDB 1EDY; (Xiao et al., 2000)], as well as for domain MG2 of hα_2M [from 'macroglobulin-like' (MG) domain, see Janssen et al., 2005; Doan and Gettins, 2007; PDB 2P9R]. It was only in 2012 that a medium-resolution (4.3 Å) crystal structure of MA-induced hα_2M (closed state) was reported (PDB 4ACQ; Marrero et al., 2012), which provided structural insight into the whole molecule, except for the very N- and C-termini and the central flexible bait region (Figures 2A, B, E, F and 3).

Protomer structure of MA-induced hα_2M

The hα_2M protomer consists of nine domains in tandem, namely MG1-MG7, CUB [name derived from C1r/C1s, urchin embryonic growth factor, and bone morphogenetic protein 1; (Bork and Beckmann, 1993)] and RBD plus two inserted domains: BRD within MG6 and TED within CUB [(Marrero et al., 2012); see Figure 2A, B and E]. This topology was generally well predicted by Doan

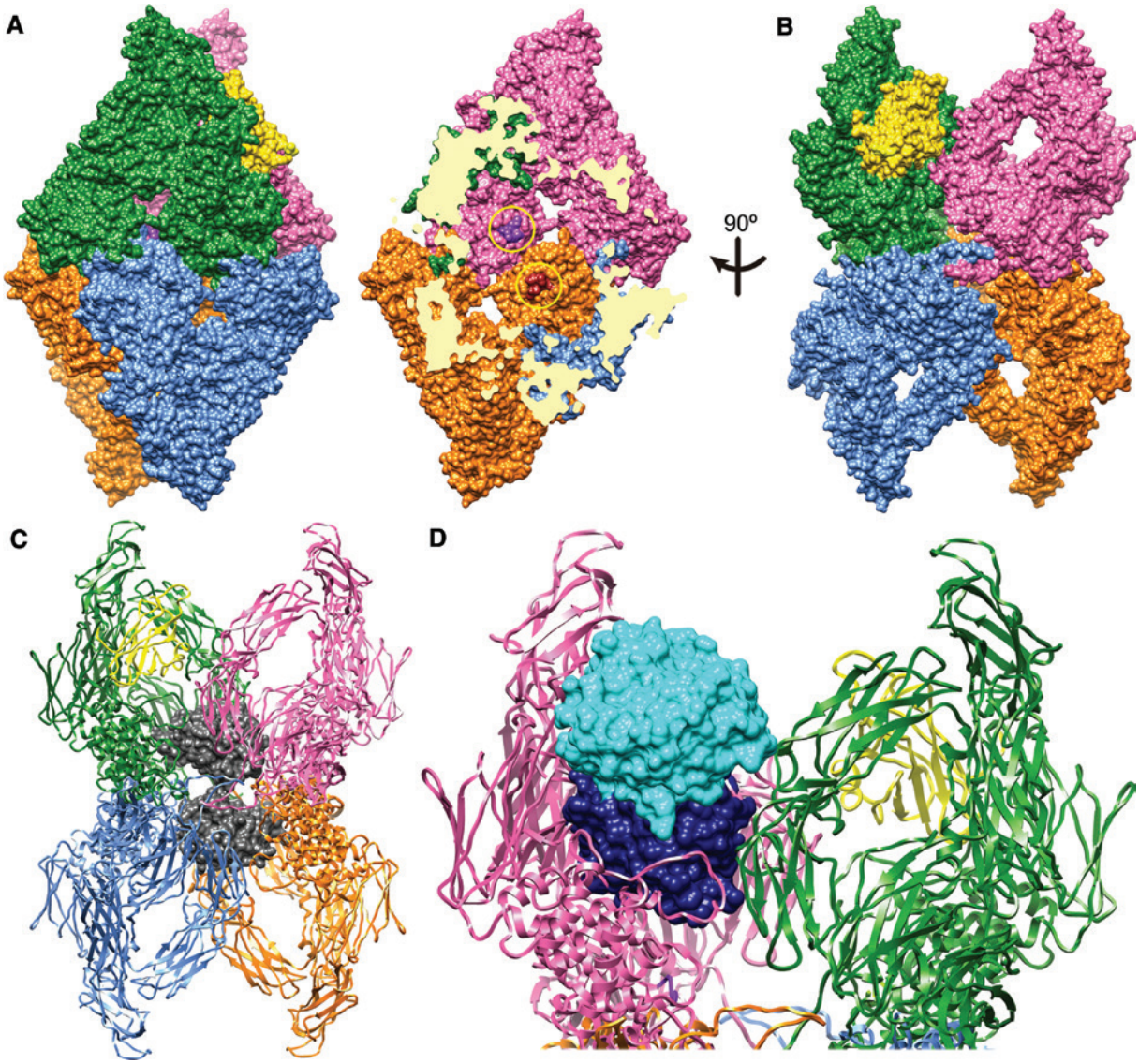

Figure 3: Tetrameric induced hα_2M.
(A) Surface representation of the MA-treated hα_2M tetramer in 'X-view' with each protomer in one color (left panel). The RBD, which is only visible in one of the protomers of the structure, is further shown in yellow. For the green protomer, the molecule in magenta is the 'disulfide-linked' protomer, the one in blue is the 'vicinal' protomer, and the one in orange is the 'opposite' protomer. Insight into the central cavity (right panel) is provided after clipping off the frontal part of the tetramer. The two thioester sites of the vicinal back monomers (magenta and orange) are encircled in yellow. (B) Tetramer in 'H-view' resulting from vertical clockwise 90°-rotation of (A). (C) Ribbon-plot representation of the tetramer in H-view modeled with two trapped molecules of HIV-1 proteinase as dark gray solid surfaces (PDB 4LL3; Kožišek et al., 2014), optimally orientated and placed to minimize clashes. (D) Suggested interaction mode of soybean trypsin Kunitz-type inhibitor with porcine trypsin (PDB 1AVW; Song and Suh, 1998), shown as cyan and dark blue solid surfaces, respectively, across an entrance 2 of the upper disulfide-linked dimer. View as in (B) and (C) after a vertical 180°-rotation.

and Gettins in 2007 in the absence of the α_2M structure, solely based on the structure of complement component C3 (Doan and Gettins, 2007). In hα_2M, the domains are arranged as a compact triangular prism, which is cross-linked by 13 disulfide bridges and has an outer convex (Figure 2E, left) and an inner concave surface (Figure 2E, right). The first seven MG domains are ~100–110-residue β-sandwiches consisting of four- and three-stranded antiparallel β-sheets (see figures 2E and 1D in Marrero et al., 2012). The sheets are rotated vertically relative to each other and joined by a central hydrophobic core, as previously described for recombinant MG2 (see the Section 'Structural aspects of tetrameric α_2M inhibitors' above, and Doan and Gettins, 2007), which conforms

to the S-type immunoglobulin-fold topology typical of fibronectin type-III domains (Chothia and Jones, 1997). Domains MG1-MG6 are arranged as a 1.5-turn MG-keyring encircling an ellipsoidal aperture of ~15 Å × ~30 Å named 'entrance 1', access to which is modulated by the protruding glycan linked to N_{247} (amino acid numbering of hα_2M in subscript following UniProt database entry [UP] P01023; deduct 23 for classical numbering of the purified protein; for glycosylation sites in hα_2M, see Figure 2A).

Inserted between the fourth and fifth strands of MG6, the BRD is an extended flexible domain that spans ~85 Å, with little regular secondary structure and few intra-domain contacts. It vertically lines the protomer on its concave face and contains three short helical segments and a short β-strand. Overall, this domain interacts with MG1, MG2, MG3, and MG5. Residue C_{689} is the last residue of the BRD defined in the crystal structure and is fixed to the protomer moiety by a disulfide-bond with upstream residue C_{642}. After C_{689}, the chain enters the flexible 39-residue bait-region (P_{690}-T_{728}; Sottrup-Jensen et al., 1981a) and the polypeptide chain is only defined again in the structure from residue E_{729} onwards, within the BRD-MG6-linker. The distance between C_{689} and E_{729} is ~90 Å, which can easily accommodate the missing residues in a continuous chain. This explains why MA-treated hα_2M, which has an intact bait-region segment (Armstrong and Quigley, 1999), can undergo a similar conformational rearrangement to the peptidase-induced form, in which the bait region is cleaved. In turn, the intrinsic disorder and flexibility of the bait region is in agreement with its function as universal, exposed bait for endopeptidases (Sottrup-Jensen et al., 1989). Moreover, C_{689} is on the concave side of the protomer, close to MG1 and MG5, while E_{729} is on the outer convex surface, close to MG7 (Figure 2E, green arrows). This indicates that the flexible bait region runs across entrance 1.

Downstream of MG6, MG7 is a hinge domain that closes the MG-keyring and leads to the 116-residue CUB domain. This is a β-sandwich consisting of two four-stranded antiparallel β-sheets (see figure 1E in Marrero et al., 2012). CUB is roughly perpendicular to MG7, which causes the former to laterally attach to domain MG2 (Figure 2E). Domain TED is inserted between the third and fourth β-strands of CUB and is a 315-residue α/α-propeller made up of 12 α-helices arranged as six concentric α-hairpins in tandem surrounding a central propeller shaft (see figure 1F in Marrero et al., 2012). A β-hairpin is inserted in the loop connecting the helices of the fifth α-hairpin. This α/α-toroid architecture is generally found in a structural superfamily named 'six-hairpin glycosidases' (Gough et al., 2001), and in hα_2M it gives rise to a thick disc with two parallel flat sides. TED is located below CUB and adjacent to MG1 and MG2, and includes the thioester bond. In native hα_2M, the latter is formed between the cysteine side-chain sulfur and the glutamine side-chain carbonyl of conserved motif C_{972}-G-E-Q_{975} (Sottrup-Jensen, 1987; Armstrong and Quigley, 1999). This bond is highly reactive so it is likely to be buried in the native structure. Upon induction through MA-treatment, the bond is cleaved and the resulting free cysteine and MA-modified glutamine are found exposed on the entry surface of TED (Figure 2E and F). Overall, the architecture of TED in hα_2M, which is generally similar to the high-resolution structure of complement fragment C3d (Nagar et al., 1998), suggests that a thioester site with equivalent geometry is also found in other thioester-containing α_2M orthologs and paralogs. After TED, the polypeptide chain completes the CUB structure and leads to the C-terminal 129-residue RBD, which lies behind and is in contact with CUB, is close to MG3 and approaches the TED β-hairpin. RBD is a variant of the MG fold with an extra β-α-β module (see figure 1G in Marrero et al., 2012). This notwithstanding, it is also known as domain 'MG8' (Doan and Gettins, 2007).

Quaternary arrangement of MA-induced hα_2M

Four structurally equivalent hα_2M protomers form the large MA-induced particle that is ~210 Å long and ~140 Å wide and deep (Figure 3A–C). This structure should also be a valid model for peptidase-induced hα_2M, as previously suggested (Barrett and Starkey, 1973; Travis and Salvesen, 1983; Sottrup-Jensen, 1989). All eight N-linked glycosylation sites of each protomer (Figure 2A) are on the surface of the particle and probably contribute to the high solubility and concentration of the protein in plasma serum (Ganrot et al., 1967). Traditionally, low-resolution structural studies suggested three possible orthogonal orientations of the particle: the 'X-view' (Figure 3A; orientation of green protomer as in Figure 2E, left), the 'H-view' (Figure 3B), and the 'End view', obtained from the H-view through a horizontal 90-degree rotation (Andersen et al., 1995). The overall structure is similar to peptidase- and MA-induced tetramers in low-resolution electron microscopy and crystallography studies (see Section 'Structural aspects of tetrameric α2M inhibitors' above). Notably, the apical protrusions observed in the H-view and originally attributed to the RBDs (see figure 4B in Kolodziej et al., 2002) actually correspond to the MG7 domains (Marrero et al., 2012).

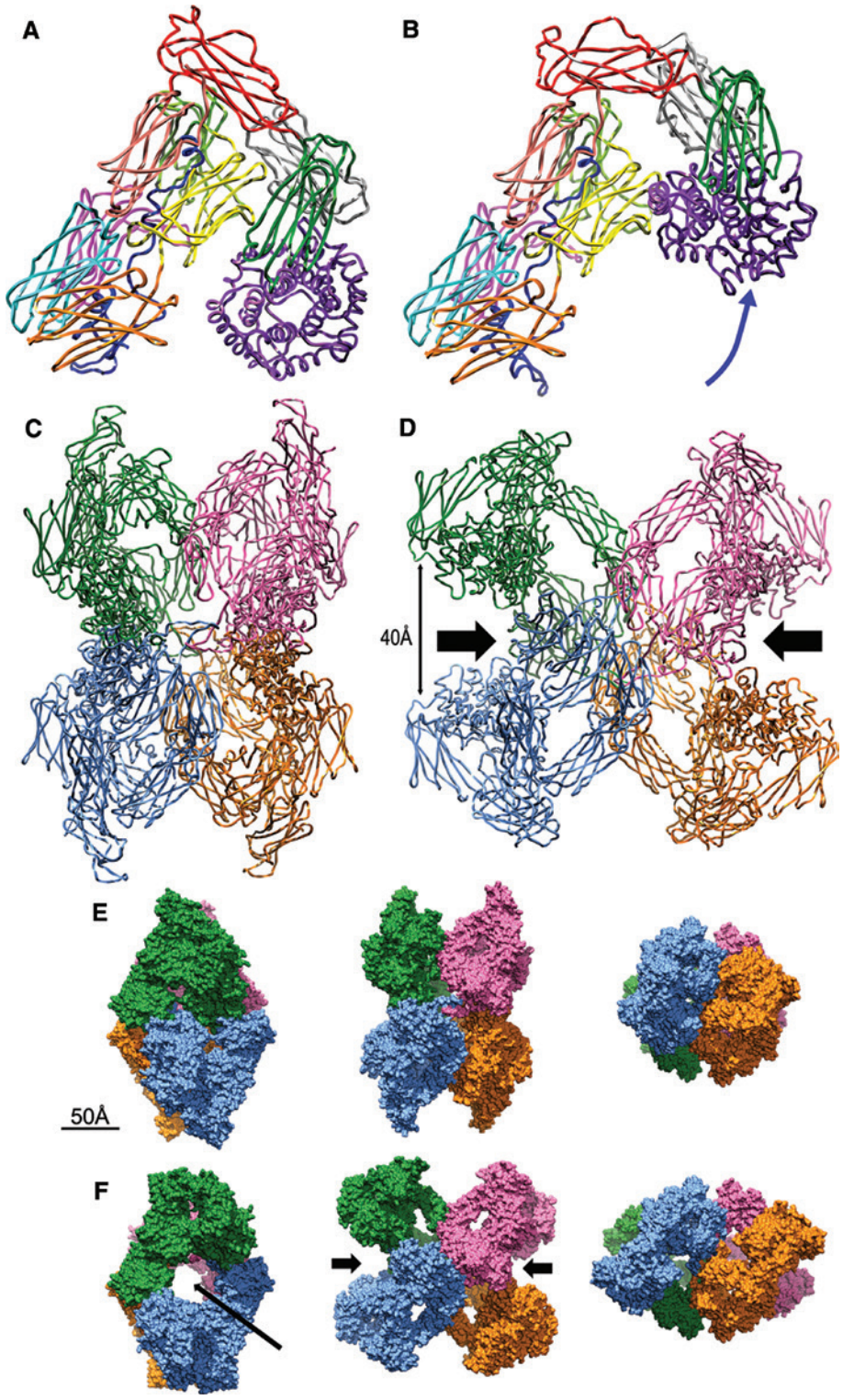

Figure 4: Hypothetical model of native hα_2M.
(A) Experimental structure of the MA-induced hα_2M protomer, and (B) computational homology model of the native hα_2M protomer based on the domain arrangement found in complement component C3 before induction (see Marrero et al., 2012). The blue arrow highlights the movement of the TED domain. (C) Experimental induced hα_2M tetramer with each protomer in one color and (D) modeled native hα_2M tetramer. Black arrows point to the potentially large entrances to the central cavity, which would span ~40 Å in diameter. (E) Space-filling model of the experimental induced tetramer structure in X-view (left), H-view (center) and End-view (right). (F) Same as (E) but for the model of native hα_2M. Black arrows point to the large entrances to the central cavity. Figure reproduced from the supplementary materials of Marrero et al. (2012), with permission from John Wiley and Sons.

Two hα_2M protomers are symmetrically linked by two disulfide bonds between domains MG3 (through C_{278}) and MG4 (through C_{431}; see Figure 2A), thus yielding a covalent dimer (protomers in green/magenta and blue/orange in Figure 3A and B). Two such dimers non-covalently associate to a tetramer, and so any protomer (e.g. the green one in Figure 3A and B) has a 'vicinal' protomer (in blue), an 'opposite' protomer (in orange), and a 'disulfide-linked' protomer (in magenta). The inter-molecular surfaces that lead to the tetramer span ~5200 Å^2. The particle is hollow and has a lumen volume of ~167,000 Å^3, which suffices to allocate two small single-domain peptidases the size of HIV proteinase or trypsin (Figure 3C) or one large peptidase the size of plasmin or factor Xa, assuming that only the catalytic domains enter the particle while ancillary domains are kept outside (Jacobsen and Sottrup-Jensen, 1993). The lumen is divided into a central 'prey chamber', which spans ~60 Å in diameter but is restrained at half height to ~30 Å diameter by a 'cavity narrowing belt'. The four thioester-bond sites are in the center of the chamber and easily accessible by prey lysines (Figure 3A, right). The prey chamber is complemented by four 'substrate ante-chambers' delimited by the concave face of each protomer (see figure 2G in Marrero et al., 2012).

Access to the central particle lumen from outside is provided by a total of 12 openings, which include entrance 1 of each protomer (see Section 'Protomer structure of MA-induced hα_2M'), as well as 'entrances 2', spanning ~30 Å × ~35 Å, and 'entrances 3', spanning ~15 Å × ~20 Å. The size of these entrances prevents caught preys from escaping the trap rapidly but enables small protein substrates or inhibitors to enter the central cavity and react with the trapped peptidase (Barrett and Starkey, 1973; Bieth et al., 1981; Travis and Salvesen, 1983). In this context, 20-kDa soybean trypsin inhibitor is the largest molecule described to interact with a prey: it slowly inhibits its hα_2M-bound porcine trypsin (Bieth et al., 1981). The size of this Kunitz-type inhibitor precludes complete intrusion through any of the 12 entrances, but modeling studies (Marrero et al., 2012) revealed that only the reactive-site loop of the inhibitor would need to be introduced through entrance 2 to efficiently bind and inhibit trypsin following the 'standard-mechanism' of inhibition (Laskowski Jr. and Kato, 1980) (Figure 3D). This hypothesis would only imply minor rearrangement of the loops shaping entrance 2, and this adjustment, in turn, would explain the slow inhibition observed. Moreover, the fact that non-covalently trapped peptidases slowly dissociate from the inhibitory complex (Wang et al., 1983) would likewise entail some rearrangement in the form of 'breathing' motions of the tetramer once induced, which would lead to enlarged entrances or even dimer dissociation.

Hypothetical model of native hα_2M and trap-closure mechanism

In the absence of medium- or high-resolution experimental structures of native tetrameric α_2Ms, a computational working model for native hα_2M was obtained by transferring the structural rearrangement between native and induced monomeric complement component C3 to the MA-induced human inhibitor (Figure 4A and B), and further applying restraints due to its tetrameric oligomerization (Figure 4C and D; Marrero et al., 2012). This homology model is similar to low-resolution electron microscopy reconstructions (Tapon-Bretaudiere et al., 1985; Delain et al., 1988; Sottrup-Jensen, 1989), and would be larger and broader than the MA-induced structure (compare panels E and F of Figure 4). This, in turn, would explain the lower electrophoretic mobility of the 'slow' native form. In addition, two large, roughly circular openings of ~40 Å in diameter would be present at the interface between vicinal protomers, on opposite sides of the tetramer, and enable access of globular peptidases of up to ~20–25 kDa molecular mass (Figure 4F, black arrows). These openings would be symmetrically shaped by domains MG1, MG2 and TED of each vicinal protomer and accessible independently, which explains why two proteinases can be trapped simultaneously (Sottrup-Jensen et al., 1981b). The model also considers that the two bait-region segments of each disulfide-linked dimer would run in parallel across the central prey chamber. Accordingly, a simple vertical rotation of the prey inside the central chamber would provide access to various bait regions, so that these could be cleaved fast in a sequential manner before conformational rearrangement occurs. This is consistent with reports stating that between two and four chains may be cleaved by a single prey endopeptidase before becoming trapped (Christensen and Sottrup-Jensen, 1983). Moreover, depending on the kinetics of the composite inhibitory reaction for a particular endopeptidase, which includes bait region cleavages, conformational changes, thioester-bond activation, and, optionally, covalent anchoring of the endopeptidase, a second prey molecule may intrude into the tetrameric particle through the second large opening (Figure 4F, center panel). Accordingly, capture of two molecules might be simultaneous rather than sequential, as previously postulated (Sottrup-Jensen, 1989; Sottrup-Jensen et al., 1989). However, under certain circumstances 'half-reacted' α_2M

forms are obtained, for example when the bait region is cleaved in only one disulfide-bridged dimer by sterically-hindered matrix-bound chymotrypsin (Qazi et al., 1998). Half-transformed molecules may subsequently bind a second proteinase molecule after cleavage of the other bait regions.

Biochemical mechanism of monomeric α_2M inhibitors

In addition to tetrameric α_2Ms, metazoans have functional monomeric forms, including α_1-inhibitor-3 (α_1I_3) from rat, guinea pig and hamster; four mouse murinoglobulins; and human α2ML1 (Enghild et al., 1989a; Rubenstein et al., 1993; Overbergh et al., 1994; Galliano et al., 2006). Monomeric α_2Ms have also been found in the serum of birds, bullfrog and a snake (Rubenstein et al., 1993). At the sequence level, monomeric α_2Ms are generally very similar to tetrameric forms, for example, rat α_1I_3 and human α2ML1 are 58% and 40% identical to hα_2M, respectively. Consistently, metazoan monomeric α_2Ms are secreted and serve as endopeptidase inhibitors in the circulation. They similarly undergo a large conformational rearrangement upon induction triggered by cleavage in a bait region that leads to exposure of the RBD. The latter is recognized by cell-surface receptors for internalization and lysosomal destruction of the inhibitor-peptidase complex (Enghild et al., 1989b; Thogersen et al., 2002). These sequence and chemical similarities led to the hypothesis that monomers may represent evolutionary precursors of the multimeric forms (Rubenstein et al., 1993; Dodds and Law, 1998).

However, considerable differences with tetrameric hα_2M are observed at the functional level as revealed by studies with rodent α_1I_3, monomeric rodent α_2M, and human α2ML1 (Gliemann and Sottrup-Jensen, 1987; Enghild et al., 1989a; Grøn et al., 1996; Galliano et al., 2006) (compare with the Section 'Biochemical mechanism of tetrameric human α_2M inhibitor'). Accordingly, in monomeric α_2Ms, peptidases cleave the bait region with or without concomitant covalent cross-linking of the peptidase, but inhibition only occurs upon cross-linking; opening of the thioester bond alone does not cause induction, so that treatment with MA only leads to induction if an endopeptidase subsequently cleaves the bait region; monomeric α_2Ms are less efficient than dimeric or tetrameric forms and require high molar excesses for significant inhibition; a fraction of monomeric inhibitor dimerizes through the association of two already rearranged induced protomers, and dimerization is not responsible for inhibition; and inhibition of bound endopeptidases is observed against very large substrates only. This may also provide an explanation for the existence of monomeric forms when the more efficient tetrameric inhibitors are constitutively abundant: they provide complementarity when only inhibition in front of very large substrates is required or may have other functions rather than peptidase inhibition (Rubenstein et al., 1993). Monomeric α_2Ms have also been studied from bacteria (bα_2Ms), where they occur in two independent forms: either with a thioester bond and represented by *E. coli* α_2M (YfhM *alias* ECAM; Doan and Gettins, 2008) and *S. typhimurium* α_2M (Sa-A2M); or co-transcribed with other proteins and without a thioester bond, such as *E. coli* YfaS and *P. aeruginosa* MagD (Budd et al., 2004; Doan and Gettins, 2008; Neves et al., 2012; Robert-Genthon et al., 2013; Wong and Dessen, 2014; Garcia-Ferrer et al., 2015). Little is known about their physiology and function in general. The best-characterized form is ECAM, whose 1,653-residue multi-modular architecture contains an N-terminal signal peptide for secretion (M^1-G^{17}; amino acid numbering of ECAM in superscript following UP P76578), a BRD and a thioester motif (see Figure 2C for ECAM domain organization and acronyms; Doan and Gettins, 2008; Garcia-Ferrer et al., 2015). ECAM and related bα_2Ms are thus ~200-residues longer than metazoan α_2Ms. Cell-fractionation experiments revealed that ECAM was found associated with the inner membrane of the periplasm of *E. coli* (Doan and Gettins, 2008; Garcia-Ferrer et al., 2015). These data, together with a report on *P. aeruginosa* MagD (Robert-Genthon et al., 2013), are consistent with ECAM and other bα_2Ms being anchored to the periplasmic side of the inner membrane through 'lipobox'-mediated lipidic linkage of the N-terminal cysteine residue of the secreted protein (C^{18} in ECAM; Doan and Gettins, 2008). In biophysical studies, native ECAM (nECAM) and MA-treated ECAM (MA-ECAM) were both monomeric and equivalent, which indicates that MA-mediated opening of the thioester bond of nECAM only causes induction if it is followed by cleavage in the bait region (Doan and Gettins, 2008; Garcia-Ferrer et al., 2015). Similar results were reported for Sa-A2M, which has 82% sequence identity with ECAM and shows essentially identical structures for the native and MA-treated inhibitor (Wong and Dessen, 2014).

Induced ECAM (iECAM) was obtained by cleavage in the bait region through endopeptidases of disparate specificity. In all cases, cleavage was efficient and showed that the bait region of ECAM is promiscuous and contains several accessible recognition sites for peptidases, despite being shorter than in hα_2M (25 residues, see Figure 2C, vs. 39 residues, see Figure 2A and Sottrup-Jensen et al.,

1989; Garcia-Ferrer et al., 2015). Protein iECAM was also monomeric but diverged in its biophysical properties from nECAM, which was attributed to large conformational rearrangement upon induction (Doan and Gettins, 2008; Neves et al., 2012; Wong and Dessen, 2014; Garcia-Ferrer et al., 2015). This is supported by the large differences observed between the structures of native Sa-A2M and iECAM (see also the Section 'Homology modeling of native ECAM and trap-closure mechanism' below) and by studies with *P. aeruginosa* MagD (Robert-Genthon et al., 2013). Collectively, these results indicate that ECAM is a pan-proteinase target protein and that cleavage in the bait region is necessary and sufficient for generating iECAM. Some peptidases also cleaved between the first and second domains of ECAM (MG0 and NIE; see Figure 2C and the Section 'Homology modeling of native ECAM and trap-closure mechanism' below), which removed the membrane anchor and thus yielded soluble iECAM through shedding. Such cleavage also caused a fraction of monomeric iECAM to be slowly transformed into dimers (Doan and Gettins, 2008; Garcia-Ferrer et al., 2015). Dimers of iECAM were stable and separable from monomers, and were structurally equivalent to iECAM monomers, which supports the idea that once induction has occurred on monomeric ECAM, dimerization just entails non-covalent protomer association without further significant rearrangement.

ECAM acted as an inhibitor of endopeptidases of disparate specificity and architecture but only against very large substrates the size of aldolase (160 kDa) or larger (Doan and Gettins, 2008; Garcia-Ferrer et al., 2015). Consistently, ECAM protected *E. coli* cell-envelope preparations, which included large murein lipoproteins, from digestion by endopeptidases. Complementary *in vivo* studies revealed that wild-type *E. coli* cells grown anaerobically had higher cell survival in the presence of peptidases than the ECAM knockout (Garcia-Ferrer et al., 2015). This is consistent with the protein being an acute-phase protein that is significantly upregulated under anaerobic conditions while hα_2M is constitutively abundant (Brokx et al., 2004). Furthermore, the inhibitory mechanism of ECAM and Sa-A2M absolutely required covalent bonding of prey peptidases by means of an intact thioester bond, which was targeted by a lysine from the prey (Doan and Gettins, 2008; Neves et al., 2012; Wong and Dessen, 2014; Garcia-Ferrer et al., 2015). Collectively, the functional and biophysical results for ECAM and related forms such as Sa-A2M and *P. aeruginosa* MagD, as well as the presence of related genes in colonizing Gram-negative bacteria, suggest a function of bα_2Ms as a bacterial cell-envelope protector against host endopeptidases (see Figure 5 and Doan and Gettins, 2008; Neves et al., 2012; Robert-Genthon et al., 2013; Wong and Dessen, 2014; Garcia-Ferrer et al., 2015).

Overall, the molecular features of bα_2Ms – but not their physiological roles – are very similar to those of monomeric mammalian α_2Ms (see above) despite disparate origins, which suggests that the overall architecture and working mechanism might be shared among monomeric α_2Ms. In general, monomers provide less efficient shields than tetrameric cages, which entails that covalent linkage through the thioester bond is indispensable for inhibition. As this linkage is rather inefficient because it relies on the presence of nearby lysines on the surface of the prey, monomeric α_2Ms are substantially less competent than their tetrameric counterparts and only cause inhibition in front of very large substrates. Thus, while probably keeping the general chemical features of their tetrameric counterparts (Rubenstein et al., 1993), monomeric family members must follow a mechanism that would be rather reminiscent of a snap-trap rather than the Venus flytrap described for tetrameric α_2Ms (Figure 1B and D). Finally, one conundrum remains to be solved: human neutrophil elastase is apparently inhibited by ECAM through covalent linkage mediated by the thioester bond (Doan and Gettins, 2008), but it lacks lysines. In contrast, this peptidase was not inhibited by monomeric α_2Ms such as rat α_1I_3 and American bullfrog α_2M (Enghild et al., 1989a; Rubenstein et al., 1993).

Structural studies of monomeric α_2M inhibitors

Low-resolution SAXS or single-particle cryo-electron microscopy studies have been reported for rat α_1-inhibitor 3 (Ikai et al., 1990), *P. aeruginosa* MagD (Robert-Genthon et al., 2013), Sa-A2M (Wong and Dessen, 2014) and ECAM (Neves et al., 2012; Wong and Dessen, 2014; Garcia-Ferrer et al., 2015). High- or medium-resolution structural data are currently available for full-length wild-type Sa-A2M, both native and MA-treated (PDB 4U48, 4U59; Wong and Dessen, 2014), as well as for point-mutant Y1175G (PDB 4U4J; Wong and Dessen, 2014). Structures have also been published for nECAM fragments spanning domains MG0-NIE-MG1 (PDB 4ZJG; Garcia-Ferrer et al., 2015), NIE-MG1 (PDB 4ZJH; Garcia-Ferrer et al., 2015) and MG7-CUB(TED)-RBD (PDB 4ZIU; Garcia-Ferrer et al., 2015); full-length trypsin-induced iECAM, i.e. lacking domain MG0 (PDB 4ZIQ; Garcia-Ferrer et al., 2015); and a fragment of elastase-induced iECAM, which is essentially identical to the trypsin-induced structure but lacks domains NIE,

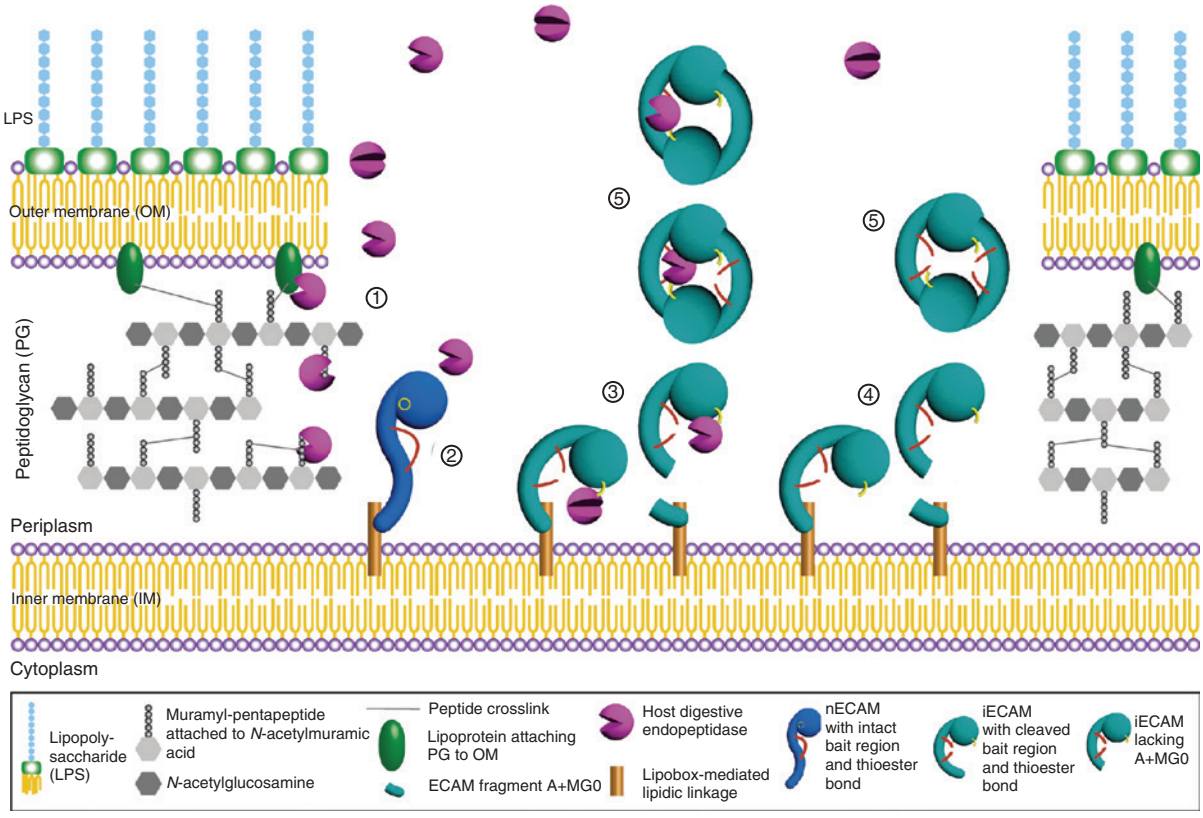

Figure 5: Proposed physiological function and working mechanism of ECAM in the periplasm of *E. coli*.
Scheme illustrating how ECAM, and possibly other bα_2Ms, may protect the bacterial peptidoglycan from extraneous endopeptidases in the gut of mammalian hosts. When the outer membrane of the bacterial cell is breached (e.g. by antibiotics, bile acids or antimicrobial peptides), host peptidases can reach the periplasm and cleave cell-wall components such as muramyl peptides and outer-membrane-anchored lipoproteins (step ①). ECAM is a large multi-modular inner-membrane broad-spectrum inhibitor of digestive host endopeptidases. In a first step, nECAM – with a hidden thioester bond and an intact bait region – is cleaved by attacking peptidases of disparate specificity (step ②). This triggers exposure of the reactive thioester bond and covalent linkage of the peptidase if any surface lysine is nearby (step ③). This process yields monomeric iECAM and entails major structural rearrangement of over half the structure (see Figure 7D), which leads to a more compact moiety. iECAM formation may occur with (step ③) or without (step ④) concomitant entrapment. In addition, membrane anchoring of ECAM ensures that only peptidases penetrating the periplasm are trapped but not those found in the extracellular milieu. Subsequently, shedding at the MG0-NIE junction yields soluble iECAM, which can be monomeric (steps ③ and ④) or, in some instances, dimeric (step ⑤). Peptidases may be trapped inside or outside iECAM dimers. Shedding enables the complex to be released to the extracellular space. There, trapped peptidases should still be active against alimentary proteins and peptides during host digestion, but not against large substrates such as bacterial cell-wall components.

MG1 and MG5 (PDB 4RTD; Fyfe et al., 2015). In addition, a homology model of full-length nECAM based on composite information provided by the nECAM fragments, single-particle cryo-electron microscopy and the coordinates of native Sa-A2M has been reported (PDB 5A24; Garcia-Ferrer et al., 2015).

Structure of trypsin-induced ECAM

The iECAM structure includes fragment P^{166}–P^{1653}, which is organized in 12 domains and a linker region (NIE to RBD; Figure 2C and D), and contains a covalently linked trypsin molecule that is disordered in the structure (see Garcia-Ferrer et al., 2015). Consistent with a bacterial protein, ECAM is not glycosylated and the only cysteines of the mature protein sequence are found in the N-terminal anchor (C^{18}; see Section 'Biochemical mechanism of monomeric α_2M inhibitors'; not present in the structure) and the thioester motif (C^{1187}, see also below), respectively. The molecule is roughly similar to hα_2M, with a front convex face (Figure 2G, left panel; reference orientation) and a back concave face (Figure 2G, right panel). However, detailed inspection reveals that the domains are arranged differently in both molecules (compare panels E and G of Figure 2).

Similarly to hα_2M, six MG domains (MG1-MG6) are arranged as a 1.5-turn MG-keyring around a central lumen of ~20 Å diameter (entrance 1), in such a way that domains MG5 and MG6 are, respectively, aligned and in contact with MG1 and MG2. As in hα_2M, MG domains are β-sandwiches, into which additional structural elements are inserted. As a result, the eight MG domains (including MG0, see the Section 'Homology modeling of native ECAM and trap-closure mechanism') span between 78 and 128 residues, i.e. they are more variable in length than in hα_2M (see Section 'Protomer structure of MA-induced hα_2M'). However, these differences do not affect the width of the domains (across the sheets; ~20 Å) or their height (between the sheets; ~15 Å), but rather their length (along the sheets; ~30–50 Å; see supplementary figure 6 in Garcia-Ferrer et al., 2015). Notably, extra domain NIE precedes MG1 and is a variant of an MG domain, into which an additional short strand has been inserted between the two last strands (see suppl. figure 6 in Garcia-Ferrer et al., 2015). Perpendicularly attached to MG3 and MG6, domain MG7 leads to the protruding C-terminal part of the molecule, which – as in hα_2M – includes domains CUB, TED and RBD (Figure 2G and suppl. figure 6 in Garcia-Ferrer et al., 2015). The CUB domain is a β-sandwich as in hα_2M except for an extra helix inserted between strands 6 and 7. TED is inserted into CUB although at a different angle than in hα_2M. ECAM TED is likewise a six-fold α/α-toroid but here the β-ribbon found in hα_2M (see Section 'Protomer structure of MA-induced hα_2M') is moved to the second α-hairpin, and an extra helix is found between the first and second hairpins. The thioester segment is a fifteen-membered thiolactone ring composed of four residues, C^{1187}-L-E-Q^{1190}, which is located at the beginning of the second helix of the first α-hairpin. In agreement with an induced peptidase-bound inhibitor, the thioester bond is broken in iECAM (Figure 2H), but formed in an nECAM fragment spanning MG7-CUB(TED)-RBD (see Section 'Homology modeling of native ECAM and trap-closure mechanism'). The thioester motif is shielded in iECAM by the aforementioned TED β-ribbon. However, while the side chain of C^{1187} is surrounded by the side chains of neighboring residues, Q^{1190} points to the bulk solvent, where the disordered trypsin molecule is bound. The C-terminal domain of iECAM, RBD, shows the same core structure as in hα_2M, but has different decoration (compare figure 1G in Marrero et al., 2012, with suppl. figure 6 in Garcia-Ferrer et al., 2015). It is likewise a variant of an MG domain, but here expanded to a six-stranded front and a five-stranded back β-sheet, whose planes are rotated away by ~40° as in MG domains. A β-ribbon and two short helices are inserted into this scaffold. Despite the topological equivalence between the RBDs in hα_2M and iECAM, the RBD function in the latter is not receptor recognition. Instead, it occupies a structurally key position to interact with TED, CUB and MG7. In addition, it stabilizes the C-terminal part of the structure protruding from the MG-keyring by contacting MG2 and MG3.

The domain that differs most between hα_2M and iECAM is BRD. In iECAM, BRD is inserted into MG6, spans 66 residues (S^{901}-N^{966}; vs. 127 residues in the human inhibitor), is irregularly folded and includes the bait region within segment Q^{934}-G^{958}. This domain plays an important role, not only in the conformational rearrangement when cleaved, but also in the stability of nECAM as shown in mutation studies (Garcia-Ferrer et al., 2015). BRD is defined in the structure for S^{901}-G^{938} (upstream of the trypsin cleavage site at R^{946}) and G^{949}–N^{966} (downstream of the cleavage site). The upstream segment is freely accessible and lines part of the concave surface of the monomer. The downstream segment starts at the interface between MG2 and CUB and interacts with MG2, MG7, CUB, and RBD (Figure 2D and G).

The oligomeric structure of iECAM found in the crystals is a dimer (see Figure 6A). In addition, studies of trypsin-induced ECAM in solution indicated that the protomer conformation of iECAM is equivalent in monomers and dimers (see Section 'Biochemical mechanism of monomeric α_2M inhibitors' above and cryo-electron microscopy studies in Garcia-Ferrer et al., 2015). Accordingly, the protomer in the dimer should be a valid model for the isolated monomeric inhibitor. The dimer is a large elongated particle of maximal dimensions ~180 Å × ~90 Å × ~80 Å, which surrounds a central prey chamber ~40 Å in diameter (Figure 6B). The particle contains five large openings: two intra-molecular entrances 1 (~20 Å in diameter, see above), two large inter-molecular entrances 2 at the dimerization interface (~60 Å × ~30 Å), and one entrance 3 (~20 Å in diameter, symmetrically created by TED and RBD of each protomer). These openings differ from the homonymous entrances of hα_2M. The two thioester segments are on the inner surface of the particle, and modeling indicated that a single trypsin molecule could be placed in the central prey chamber upon slight rearrangement, where it would be accessible to substrates through any of the five entrances. However, the thioesters are also close to entrance 2, so the bound prey could likewise be placed outside the cage while remaining covalently linked to the particle (see Figure 5, ⑤). Both situations would be compatible with crystal packing and would explain the activity of bound peptidases, except against very large substrates (see Section

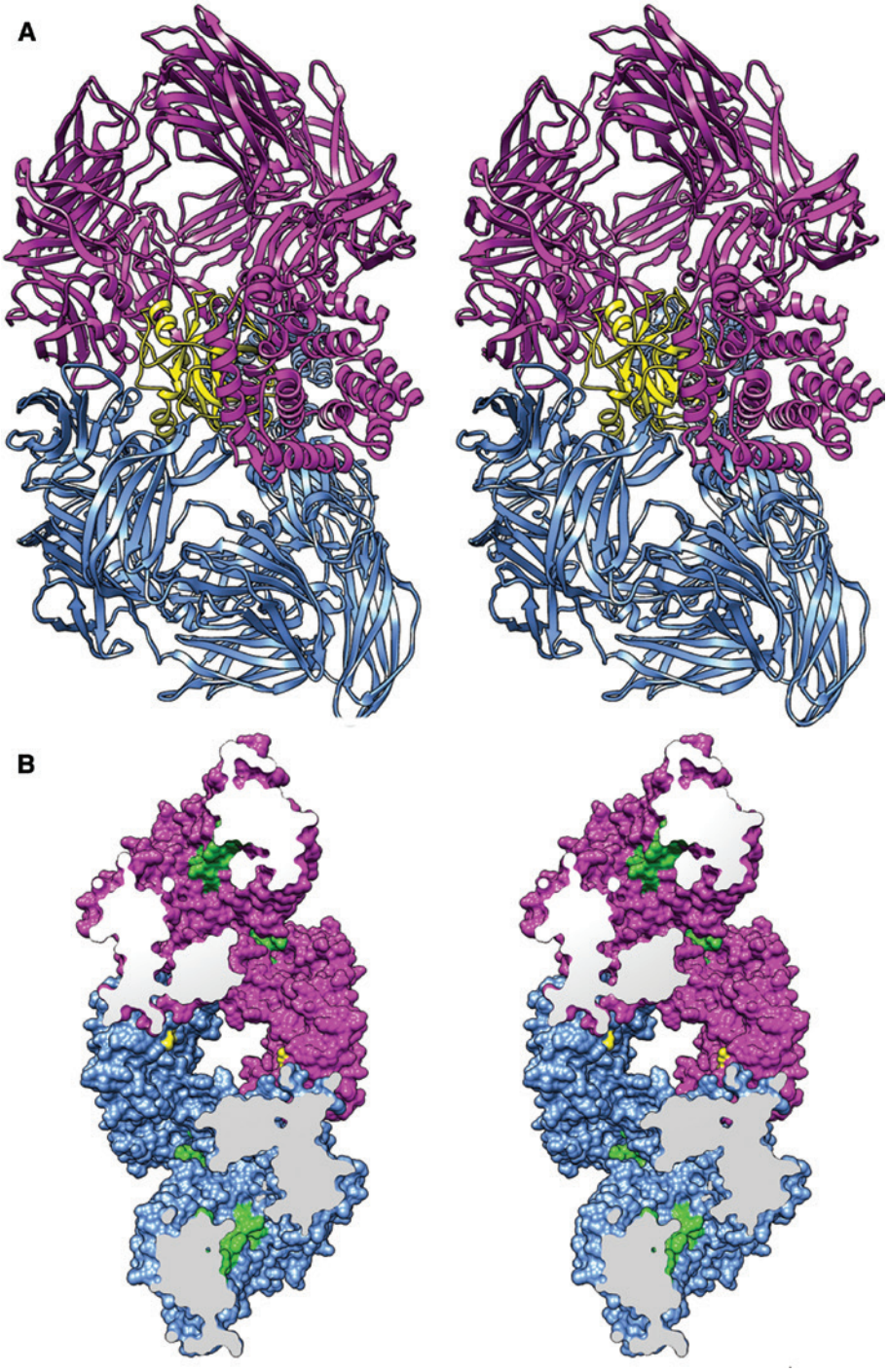

Figure 6: Dimeric iECAM structure.
(A) iECAM dimer with monomers in magenta and blue. The depicted front view is perpendicular to the crystallographic two-fold relating both monomers. A trypsin molecule (yellow ribbon) has been tentatively placed in the central prey chamber for reference. (B) View of a cut-through surface model of the dimer in stereo colored as in panel (A), which provides insight into the central prey chamber. Thioester residues in yellow; bait-region domains in green.

'Biochemical mechanism of monomeric α_2M inhibitors'). In addition, as reported for similar cage structures, the dimer may undergo 'breathing' motions of segments surrounding the trapped prey as suggested for hα_2M (Marrero et al., 2012). However, as in monomeric mammalian α_2Ms (see Section 'Biochemical mechanism of monomeric α_2M inhibitors'), dimerization is not a universal mechanism for ECAM. Instead, it is restricted to

specific peptidases, so that the induced monomer is the most abundant species.

Homology modeling of native ECAM and trap-closure mechanism

In the absence of a crystallographic structure of full-length nECAM, a model was constructed based on composite experimental information provided by nECAM fragments spanning MG0-NIE-MG1, NIE-MG1 and MG7-CUB(TED)-RBD; cryo-electron microscopy studies and the native coordinates of the close relative Sa-A2M (Wong and Dessen, 2014; Garcia-Ferrer et al., 2015). The resulting nECAM homology model (MG0-to-RBD; K^{57}-P^{1653}) revealed an extended helicoidal moiety of ~160 Å length (Figure 7A). MG0, at the N-terminus of ECAM, is expected to be close to the inner membrane *in vivo* and flexibly linked with NIE, so it is easily removed during induction and release of soluble iECAM (see Section 'Biochemical mechanism of monomeric α_2M inhibitors' and Figure 5). As in iECAM, nECAM would feature a central MG-keyring but rearranged so that it nearly lacks a central entrance 1. Segments MG0 and MG7-CUB(TED)-RBD would protrude from the MG-keyring in opposite directions. BRD, in turn, would include three segments in helical conformation, as in iECAM, and line the concave surface of the MG-keyring (Figure 7B), thus interacting with linker L and domains MG1, MG2, and MG4-MG6. The bait region would be flexible and freely accessible for prey peptidases. In addition, the thioester region within TED is buried in nECAM, as indicated by the experimental MG7-CUB(TED)-RBD structure, and faces the outer surface of the six-stranded front sheet of RBD (Figure 7C). The thioester bond itself is intact (Figure 7C and E) and protected by TED and RBD. Involvement of the latter in this protection underpins its importance in overall ECAM structural integrity and stability (see also the Section 'Structure of trypsin-induced ECAM' and below).

A comparison of iECAM and nECAM reveals the likely molecular mechanism of ECAM induction, which entails conformational rearrangement that gives rise to a more compact structure (Figure 7D). This is consistent with differences in the biophysical properties of both species (Garcia-Ferrer et al., 2015). The structures only coincide on the bilayered side of the MG-keyring – MG1-L-MG2 and MG5-MG6 – and, partially, at BRD (up to Y^{932} and from H^{964} onwards). Upon induction, MG3 and MG4 are flipped inward towards MG6 as a rigid body, due to a ~90° rotation around the anchor point of MG3 with MG2 and a simultaneous translation downward of ~50 Å. The new position of MG4 drives NIE out along the outer surface of the four-stranded sheet of MG1. This rearrangement traps the segment of the bait region upstream of the cleavage site following a ~180° rotation downward around G^{933}. Downstream in the sequence, the bait region is undefined from Q^{939} to G^{948} in iECAM and the distance between the flanking residues (66 Å) is too great to be covered by the ten missing residues. In contrast, the corresponding distance between bait-region anchor points can be easily covered by the missing residues in induced hα_2M (see Section 'Protomer structure of MA-induced hα_2M' above). This explains why the bait region must be cleaved in ECAM for induction, while in hα_2M an intact bait region is compatible with the induced form, so that it can be obtained by MA-treatment only (Marrero et al., 2012).

The rearrangement of MG3 is simultaneous with MG7 and RBD becoming rotated as a rigid body by ~25° downward, so that RBD is displaced by ~25 Å towards MG3. Rearrangement of MG7 and RBD also causes CUB and TED to be rotated and translated: the former by ~35 Å and the latter by ~45 Å. When only these two domains are compared, CUB is rotated by ~90° with respect to TED around the domain interface. These changes also cause the first helix of TED to unwind and segment Y^{1183}-G^{1186}, which acts as a protective lid of the thioester bond in nECAM, to be displaced (Figure 7E). In this way, the thioester bond becomes exposed and solvent-accessible in iECAM, so that it can be targeted by lysines. Interestingly, the initial movement of MG3 relative to MG2 is suppressed in nECAM by the BRD segment after the bait region, which passes above the MG2-MG3 linker (Figure 7B). Upon cleavage in the bait region, this constraint is released and the segment downstream of the cleavage site adopts its induced position on the outer surface of CUB.

Taken together, these structural details back the functional characterization of ECAM *in vitro* (see Section 'Biochemical mechanism of monomeric α_2M inhibitors') and indicate that the inhibitor functions as an irreversible snap trap, similarly to the otherwise unrelated serpin inhibitors (Huntington, 2013). In the latter, cleavage of a reactive-site bond imbedded in a reactive-center loop through serine and cysteine endopeptidases forms a covalent (thio)acyl-enzyme intermediate. This causes rapid and very large conformational rearrangement of the serpin and the free N-terminal part of the reactive-center loop to insert as a new strand into a central β-sheet of the inhibitor moiety. This occurs under translocation of the covalently attached peptidase to the distal side of the inhibitor moiety (Gettins, 2002). Conceptually similar but functionally dissimilar, monomeric nECAM represents the baited and set trap, with a spring-loaded bar (the hidden thioester bond) and a trip (the BRD segment after the bait region) to release

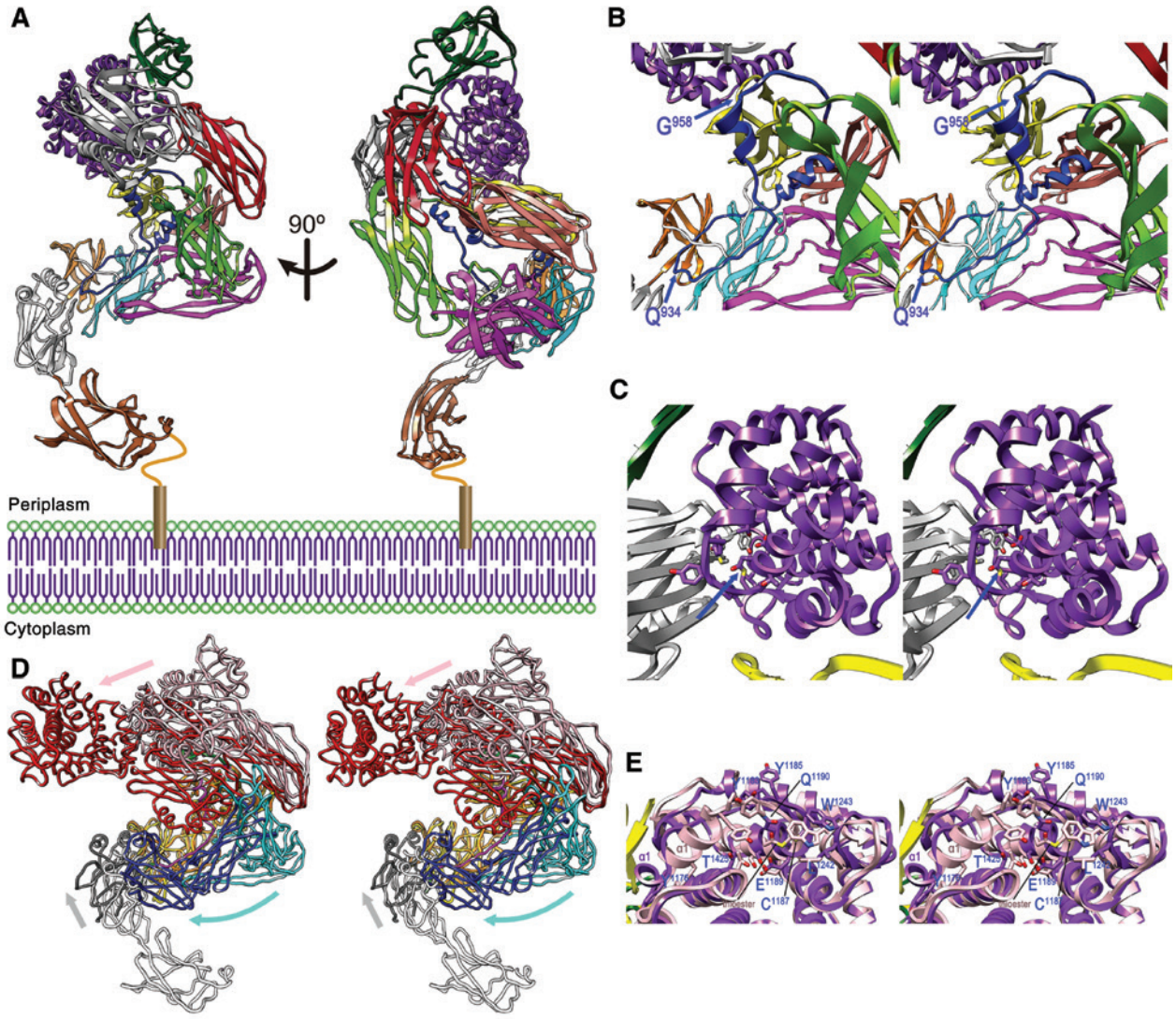

Figure 7: Homology model of native ECAM and structure-based mechanism of induction.
(A) Composite model of full-length nECAM in two orthogonal views; the domains are colored as in Figure 2C and G. (B) Close-up of (A) showing the BRD with the suggested limits of the bait region (blue arrows). (C) Close-up of (A) centered on the TED-RBD interface, which occludes the intact thioester [blue arrow; see also (E)]. (D) Model of nECAM [as in (A)] and structure of iECAM in stereo after optimal superposition of MG1-L-MG2 and MG5-MG6 (both in orange for iECAM, in yellow for nECAM). Significantly rearranged moieties are MG7-CUB(TED)-RBD (iECAM, red; nECAM, pink); MG3-MG4 (iECAM, dark blue; nECAM, cyan); and MG0-NIE (iECAM, – NIE only –, gray; nECAM, white), whose rearrangement is illustrated by arrows. (E) Close-up in stereo showing TED centered on the thioester region in both iECAM (purple) and nECAM (pink). Some residues of iECAM are labeled for reference, as is the intact thioester bond of nECAM.

it (Figure 1D). When the bait region is cleaved, large conformational rearrangement leads to induction and exposure of the thioester bond, which is analogous to firing the trap through the rapid swing-down of the spring-loaded bar. However, the prey is only trapped if covalently bound by the thioester bond. Independently of the prey binding, the trap remains irreversibly inactivated once triggered. However, contrary to a true snap trap the prey peptidase is not completely disabled by ECAM, but merely restricted in its radius of action and substrate size. Finally, the many similarities in function and biochemistry among bacterial and animal monomeric α_2M allow us to suggest that the latter probably function following similar structural pathways to their bacterial counterparts.

Outlook

Metazoan α_2Ms and some bacterial orthologs, which may be xenologs acquired through horizontal gene transfer

(Budd et al., 2004), provide a unique mechanism to control proteolysis. The mechanism does not entail inactivation or blocking of the active site of an endopeptidase, but rather steric shielding, covalent binding, and subsequent clearance from the site of action (Sottrup-Jensen, 1989; Doan and Gettins, 2008). This review describes recent advances in the structural characterization of tetrameric and monomeric α_2Ms, which, owing to their intrinsic differences in shielding potential, display variable inhibitory efficiencies and follow structurally divergent Venus-flytrap or snap-trap mechanisms, respectively. Further studies are required to fit more pieces into the structural jigsaw puzzle, in particular to add the high-resolution structures of native tetrameric inhibitors and native and induced dimeric forms.

Acknowledgments: This study was supported in part by grants from European, Spanish, and Catalan agencies (grant references FP7-HEALTH-2012-306029-2 'TRIGGER', BFU2015-64487R, MDM-2014-0435, BIO2013-49320-EXP, and 2014SGR9). The Structural Biology Unit (www.sbu.csic.es) of IBMB has been named a 'María de Maeztu' Unit of Excellence by the Spanish Ministry of Economy and Competitiveness. T.G. and I.G.-F. acknowledge a '*Juan-de-la-Cierva*' contract from the Spanish Ministry of Economy and Competitiveness (JCI-2012-13573) and an FPU fellowship from the Spanish Ministry of Education and Culture (AP-2010-03799), respectively.

References

Andersen, G.R., Jacobsen, L., Thirup, S., Nyborg, J., and Sottrup-Jensen, L. (1991). Crystallization and preliminary X-ray analysis of methylamine-treated α_2-macroglobulin and 3 α_2-macroglobulin-proteinase complexes. FEBS Lett. *292*, 267–270.

Andersen, G.R., Koch, T., Sorensen, A.H., Thirup, S., Nyborg, J., Dolmer, K., Jacobsen, L., and Sottrup-Jensen, L. (1994). Crystallization of proteins of the α_2-macroglobulin superfamily. Ann. N. Y. Acad. Sci. *737*, 444–446.

Andersen, G.R., Koch, T.J., Dolmer, K., Sottrup-Jensen, L., and Nyborg, J. (1995). Low resolution X-ray structure of human methylamine-treated α_2-macroglobulin. J. Biol. Chem. *270*, 25133–25141.

Andersen, O.M., Christensen, P.A., Christensen, L.L., Jacobsen, C., Moestrup, S.K., Etzerodt, M., and Thøgersen, H.C. (2000). Specific binding of α-macroglobulin to complement-type repeat CR4 of the low-density lipoprotein receptor-related protein. Biochemistry *39*, 10627–10633.

Armstrong, P.B. (2006). Proteases and protease inhibitors: a balance of activities in host-pathogen interaction. Immunobiology *211*, 263–281.

Armstrong, P.B. and Quigley, J.P. (1999). α_2-Macroglobulin: an evolutionarily conserved arm of the innate immune system. Dev. Comp. Immunol. *23*, 375–390.

Arolas, J.L., Botelho, T.O., Vilcinskas, A., and Gomis-Rüth, F.X. (2011). Structural evidence for standard-mechanism inhibition in metallopeptidases from a complex poised to resynthesize a peptide bond. Angew. Chem. Int. Ed. *50*, 10357–10360.

Barrett, A.J. (1981). α_2-Macroglobulin. Methods Enzymol. *80* (Pt C), 737–754.

Barrett, A.J. and Starkey, P.M. (1973). The interaction of α_2-macroglobulin with proteinases. Characteristics and specificity of the reaction, and a hypothesis concerning its molecular mechanism. Biochem. J. *133*, 709–724.

Barrett, A.J., Brown, M.A., and Sayers, C.A. (1979). The electrophoretically 'slow' and 'fast' forms of the α_2-macroglobulin molecule. Biochem. J. *181*, 401–418.

Bender, R.C. and Bayne, C.J. (1996). Purification and characterization of a tetrameric α-macroglobulin proteinase inhibitor from the gastropod mollusc *Biomphalaria glabrata*. Biochem. J. *316*, 893–900.

Bieth, J.G., Tourbez-Perrin, M., and Pochon, F. (1981). Inhibition of α_2-macroglobulin-bound trypsin by soybean trypsin inhibitor. J. Biol. Chem. *256*, 7954–7957.

Björk, I. and Fish, W.W. (1982). Evidence for similar conformational changes in α_2-macroglobulin on reaction with primary amines or proteolytic enzymes. Biochem. J. *207*, 347–356.

Blandin, S. and Levashina, E.A. (2004). Thioester-containing proteins and insect immunity. Mol. Immunol. *40*, 903–908.

Bode, W. and Huber, R. (1992). Natural protein proteinase inhibitors and their interaction with proteinases. Eur. J. Biochem. *204*, 433–451.

Boisset, N., Penczek, P., Pochon, F., Frank, J., and Lamy, J. (1993). Three-dimensional architecture of human α_2-macroglobulin transformed with methylamine. J. Mol. Biol. *232*, 522–529.

Boisset, N., Taveau, J.C., Pochon, F., and Lamy, J. (1996). Similar architectures of native and transformed human α_2-macroglobulin suggest the transformation mechanism. J. Biol. Chem. *271*, 25762–25769.

Bork, P. and Beckmann, G. (1993). The CUB domain. A widespread module in developmentally regulated proteins. J. Mol. Biol. *231*, 539–545.

Brokx, S.J., Ellison, M., Locke, T., Bottorff, D., Frost, L., and Weiner, J.H. (2004). Genome-wide analysis of lipoprotein expression in *Escherichia coli* MG1655. J. Bacteriol. *186*, 3254–3258.

Brown, R.K., Baker, W.H., Peterkofsky, A., and Kauffman, D.L. (1954). Crystallization and properties of a glycoprotein isolated from human plasma. J. Am. Chem. Soc. *76*, 4244–4245.

Budd, A., Blandin, S., Levashina, E.A., and Gibson, T.J. (2004). Bacterial α_2-macroglobulins: colonization factors acquired by horizontal gene transfer from the metazoan genome? Genome Biol. *5*, R38.

Chaikeeratisak, V., Somboonwiwat, K., and Tassanakajon, A. (2012). Shrimp α_2-macroglobulin prevents the bacterial escape by inhibiting fibrinolysis of blood clots. PLoS One *7*, e47384.

Chothia, C. and Jones, E.Y. (1997). The molecular structure of cell adhesion molecules. Annu. Rev. Biochem. *66*, 823–862.

Christensen, U. and Sottrup-Jensen, L. (1983). Enzymatic properties of α_2-macroglobulin-proteinase complexes. Apparent discrimination between covalently and noncovalently bound trypsin by reaction with soybean trypsin inhibitor. Biochim. Biophys. Acta *747*, 263–275.

Cohn, E.J., Strong, L.E., Hughes, W.L., Mulford, D.L., Ashworth, J.N., Melin, M., and Taylor, H.L. (1946). Preparation and properties

of serum and plasma proteins. IV. A system for the separation into fractions of the protein and lipoprotein components of biological tissues and fluids. J. Am. Chem. Soc. *68*, 459–475.

Conrad, H.E. (1998). Heparin-binding Proteins (San Diego, CA: Academic Press).

Debanne, M.T., Bell, R., and Dolovich, J. (1976). Characteristics of the macrophage uptake of proteinase-α-macroglobulin complexes. Biochim. Biophys. Acta. *428*, 466–475.

Delain, E., Barray, M., Tapon-Bretaudiere, J., Pochon, F., Marynen, P., Cassiman, J.J., Van den Berghe, H., and Van Leuven, F. (1988). The molecular organization of human α_2-macroglobulin. An immunoelectron microscopic study with monoclonal antibodies. J. Biol. Chem. *263*, 2981–2989.

Delain, E., Pochon, F., Barray, M., and van Leuven, F. (1992). Ultrastructure of α_2-macroglobulins. Electron Microsc. rev. *5*, 231–281.

Delain, E., Barray, M., Pochon, F., Gliemann, J., and Moestrup, S.K. (1994). Electron microscopic visualization of the human α_2-macroglobulin receptor and its interaction with α_2-macroglobulin/chymotrypsin complex. Ann. N. Y. Acad. Sci. *737*, 202–211.

Doan, N. and Gettins, P.G.W. (2007). Human α_2-macroglobulin is composed of multiple domains, as predicted by homology with complement component C3. Biochem. J. *407*, 23–30.

Doan, N. and Gettins, P.G.W. (2008). α-Macroglobulins are present in some Gram-negative bacteria: characterization of the α_2-macroglobulin from *Escherichia coli*. J. Biol. Chem. *283*, 28747–28756.

Dodds, A.W. and Law, S.K. (1998). The phylogeny and evolution of the thioester bond-containing proteins C3, C4 and α_2-macroglobulin. Immunol. Rev. *166*, 15–26.

Dubin, G., Koziel, J., Pyrc, K., Wladyka, B., and Potempa, J. (2013). Bacterial proteases in disease - role in intracellular survival, evasion of coagulation/ fibrinolysis innate defenses, toxicoses and viral infections. Curr. Pharm. Des. *19*, 1090–1113.

Enghild, J.J., Salvesen, G., Thøgersen, I.B., and Pizzo, S.V. (1989a). Proteinase binding and inhibition by the monomeric α-macroglobulin rat α_1-inhibitor-3. J. Biol. Chem. *264*, 11428–11435.

Enghild, J.J., Thøgersen, I.B., Roche, P.A., and Pizzo, S.V. (1989b). A conserved region in α-macroglobulins participates in binding to the mammalian α-macroglobulin receptor. Biochemistry. *28*, 1406–1412.

Fyfe, C.D., Grinter, R., Josts, I., Mosbahi, K., Roszak, A.W., Cogdell, R.J., Wall, D.M., Burchmore, R.J., Byron, O., and Walker, D. (2015). Structure of protease-cleaved *Escherichia coli* α_2-macroglobulin reveals a putative mechanism of conformational activation for protease entrapment. Acta Crystallogr. D *71*, 1478–1486.

Galliano, M.F., Toulza, E., Gallinaro, H., Jonca, N., Ishida-Yamamoto, A., Serre, G., and Guerrin, M. (2006). A novel protease inhibitor of the α_2-macroglobulin family expressed in the human epidermis. J. Biol. Chem. *281*, 5780–5789.

Ganrot, P.O., Gydell, K., and Ekelund, H. (1967). Serum concentration of α_2-macroglobulin, haptoglobin and α_1-antitrypsin in diabetes mellitus. Acta Endocrinol. (Copenhagen). *55*, 537–544.

Ganrot, P.O. and Schersten, B. (1967). Serum α_2-macroglobulin concentration and its variation with age and sex. Clin. Chim. Acta *15*, 113–120.

Garcia-Ferrer, I., Arêde, P., Gómez-Blanco, J., Luque, D., Duquerroy, S., Castón, J.R., Goulas, T., and Gomis-Rüth, F.X. (2015). Structural and functional insights into *Escherichia coli* α_2-macroglobulin endopeptidase snap-trap inhibition. Proc. Natl. Acad. Sci. USA *112*, 8290–8295.

Gettins, P.G.W. (2002). Serpin structure, mechanism, and function. Chem. Rev. *102*, 4751–4804.

Gliemann, J. and Sottrup-Jensen, L. (1987). Rat plasma α_1-inhibitor$_3$ binds to receptors for α_2-macroglobulin. FEBS Lett. *221*, 55–60.

Gough, J., Karplus, K., Hughey, R., and Chothia, C. (2001). Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. J. Mol. Biol. *313*, 903–919.

Goulas, T., Garcia-Ferrer, I., García-Piqué, S., Sottrup-Jensen, L., and Gomis-Rüth, F.X. (2014). Crystallization and preliminary X-ray diffraction analysis of eukaryotic α_2 -macroglobulin family members modified by methylamine, proteases and glycosidases. Mol. Oral Microbiol. *29*, 354–364.

Grøn, H., Thøgersen, I.B., Enghild, J.J., and Pizzo, S.V. (1996). Structural and functional analysis of the spontaneous re-formation of the thiol ester bond in human α_2-macroglobulin, rat α_1-inhibitor-3 and chemically modified derivatives. Biochem. J. *318*, 539–545.

Huang, W., Dolmer, K., Liao, X., and Gettins, P.G.W. (2000). NMR solution structure of the receptor binding domain of human α_2-macroglobulin. J. Biol. Chem. *275*, 1089–1094.

Huntington, J.A. (2013). Thrombin inhibition by the serpins. J. Thromb. Haemost. *11* (Suppl. 1), 254–264.

Iijima, M., Tobe, T., Sakamoto, T., and Tomita, M. (1984). Biosynthesis of the internal thioester bond of the third component of complement. J. Biochem. *96*, 1539–1546.

Ikai, A., Nishigai, M., Saito, A., Sinohara, H., Muto, Y., and Arata, Y. (1990). Electron microscopic demonstration of a common structural motif in human complement factor C3 and rat α_1-inhibitor 3 (murinoglobulin). FEBS Lett. *260*, 291–293.

Ikai, A., Ookata, K., Shimizu, M., Nakamichi, N., Ito, M., and Matsumura, T. (1999). A recombinant bait region mutant of human α_2-macroglobulin exhibiting an altered proteinase-inhibiting spectrum. Cytotechnology *31*, 53–60.

Imber, M.J. and Pizzo, S.V. (1981). Clearance and binding of two electrophoretic 'fast' forms of human α_2-macroglobulin. J. Biol. Chem. *256*, 8134–8139.

Jacobsen, L. and Sottrup-Jensen, L. (1993). Localization of ε-lysyl-γ-glutamyl cross-links in α_2-macroglobulin-plasmin complex. Biochemistry *32*, 120–126.

Janssen, B.J.C., Huizinga, E.G., Raaijmakers, H.C.A., Roos, A., Daha, M.R., Nilsson-Ekdahl, K., Nilsson, B., and Gros, P. (2005). Structures of complement component C3 provide insights into the function and evolution of immunity. Nature *437*, 505–511.

Jenner, L., Husted, L., Thirup, S., Sottrup-Jensen, L., and Nyborg, J. (1998). Crystal structure of the receptor-binding domain of α_2-macroglobulin. Structure *6*, 595–604.

Kantyka, T., Rawlings, N.D., and Potempa, J. (2010). Prokaryote-derived protein inhibitors of peptidases: A sketchy occurrence and mostly unknown function. Biochimie *92*, 1644–1656.

Kolodziej, S.J., Schroeter, J.P., Strickland, D.K., and Stoops, J.K. (1996). The novel three-dimensional structure of native human α_2-macroglobulin and comparisons with the structure of the methylamine derivative. J. Struct. Biol. *116*, 366–376.

Kolodziej, S.J., Wagenknecht, T., Strickland, D.K., and Stoops, J.K. (2002). The three-dimensional structure of the human α_2-macroglobulin dimer reveals its structural organization in the tetrameric native and chymotrypsin α_2-macroglobulin complexes. J. Biol. Chem. *277*, 28031–28037.

Kožišek, M., Lepšik, M., Grantz Šašková, K., Brynda, J., Konvalinka, J., and Rezáčová, P. (2014). Thermodynamic and structural analysis of HIV protease resistance to darunavir – analysis of heavily mutated patient-derived HIV-1 proteases. FEBS J. *281*, 1834–1847.

Laskowski, Jr. M. and Kato, I. (1980). Protein inhibitors of proteinases. Annu. Rev. Biochem. *49*, 593–626.

Laskowski, Jr. M. and Qasim, M.A. (2000). What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? Biochim. Biophys. Acta. *1477*, 324–337.

Marrero, A., Duquerroy, S., Trapani, S., Goulas, T., Guevara, T., Andersen, G.R., Navaza, J., Sottrup-Jensen, L., and Gomis-Rüth, F.X. (2012). The crystal structure of human α_2-macroglobulin reveals a unique molecular cage. Angew. Chem. Int. Ed. *51*, 3340–3344.

Nagar, B., Jones, R.G., Diefenbach, R.J., Isenman, D.E., and Rini, J.M. (1998). X-ray crystal structure of C3d: a C3 fragment and ligand for complement receptor 2. Science *280*, 1277–1281.

Nagase, H. and Harris, Jr. E.D. (1983). Ovostatin: a novel proteinase inhibitor from chicken egg white. II. Mechanism of inhibition studied with collagenase and thermolysin. J. Biol. Chem. *258*, 7490–7498.

Nagase, H., Harris, Jr. E.D., Woessner, Jr. J.F., and Brew, K. (1983). Ovostatin: a novel proteinase inhibitor from chicken egg white. I. Purification, physicochemical properties, and tissue distribution of ovostatin. J. Biol. Chem. *258*, 7481–7489.

Neurath, H. and Walsh, K.A. (1976). Role of proteolytic enzymes in biological regulation. Proc. Natl. Acad. Sci. USA *73*, 3825–3832.

Neves, D., Estrozi, L.F., Job, V., Gabel, F., Schoehn, G., and Dessen, A. (2012). Conformational states of a bacterial α_2-macroglobulin resemble those of human complement C3. PLoS One *7*, e35384.

Overbergh, L., Hilliker, C., Lorent, K., Van Leuven, F., and Van den Berghe, H. (1994). Identification of four genes coding for isoforms of murinoglobulin, the monomeric mouse α_2-macroglobulin: characterization of the exons coding for the bait region. Genomics *22*, 530–539.

Qazi, U., Gettins, P.G., and Stoops, J.K. (1998). On the structural changes of native human α_2-macroglobulin upon proteinase entrapment. Three-dimensional structure of the half-transformed molecule. J. Biol. Chem. *273*, 8987–8993.

Qazi, U., Gettins, P.G., Strickland, D.K., and Stoops, J.K. (1999). Structural details of proteinase entrapment by human α_2-macroglobulin emerge from three-dimensional reconstructions of Fab labeled native, half-transformed, and transformed molecules. J. Biol. Chem. *274*, 8137–8142.

Rawlings, N.D., Barrett, A.J., and Finn, R. (2016). Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. *44*, D343–D350.

Rehman, A.A., Ahsan, H., and Khan, F.H. (2013). α_2-Macroglobulin: a physiological guardian. J. Cell. Physiol. *228*, 1665–1675.

Robert-Genthon, M., Casabona, M.G., Neves, D., Coute, Y., Ciceron, F., Elsen, S., Dessen, A., and Attree, I. (2013). Unique features of a *Pseudomonas aeruginosa* α_2-macroglobulin homolog. MBio. *4*, e00309–13.

Rubenstein, D.S., Thøgersen, I.B., Pizzo, S.V., and Enghild, J.J. (1993). Identification of monomeric α-macroglobulin proteinase inhibitors in birds, reptiles, amphibians and mammals, and purification and characterization of a monomeric α-macroglobulin proteinase inhibitor from the American bullfrog *Rana catesbeiana*. Biochem. J. *290*, 85–95.

Salvesen, G.S., Sayers, C.A., and Barrett, A.J. (1981). Further characterization of the covalent linking reaction of α2-macroglobulin. Biochem. J. *195*, 453–461.

Sand, O., Folkersen, J., Westergaard, J.G., and Sottrup-Jensen, L. (1985). Characterization of human pregnancy zone protein. Comparison with human α_2-macroglobulin. J. Biol. Chem. *260*, 15723–15735.

Schultze, H.E., Göllner, I., Heide, K., Schönenberger, M., and Schwick, G. (1955). Zur Kenntnis der α-Globuline des menschlichen Normalserums. Z. Naturforsch. *10b*, 463–473.

Song, H.K. and Suh, S.W. (1998). Kunitz-type soybean trypsin inhibitor revisited: refined structure of its complex with porcine trypsin reveals an insight into the interaction between a homologous inhibitor from *Erythrina caffra* and tissue-type plasminogen activator. J. Mol. Biol. *275*, 347–363.

Sottrup-Jensen, L., Petersen, T.E., and Magnusson, S. (1980). A thiol-ester in α_2-macroglobulin cleaved during proteinase complex formation. FEBS Lett. *121*, 275–279.

Sottrup-Jensen, L., Lønblad, P.B., Stepanik, T.M., Petersen, T.E., Magnusson, S., and Jörnvall, H. (1981a). Primary structure of the 'bait' region for proteinases in α_2-macroglobulin. Nature of the complex. FEBS Lett. *127*, 167–173.

Sottrup-Jensen, L., Petersen, T.E., and Magnusson, S. (1981b). Trypsin-induced activation of the thiol esters in α2-macroglobulin generates a short-lived intermediate ('nascent' α_2-M) that can react rapidly to incorporate not only methylamine or putrescine but also proteins lacking proteinase activity. FEBS Lett. *128*, 123–126.

Sottrup-Jensen, L. (1987). α_2-Macroglobulin and related thiol ester plasma proteins. In: The plasma proteins. Structure, function and genetic control., Vol. 5, 2nd edn, F.W. Putnam, ed. (Orlando, FL: Academic Press), pp. 191–291.

Sottrup-Jensen, L. (1989). α-Macroglobulins: structure, shape, and mechanism of proteinase complex formation. J. Biol. Chem. *264*, 11539–11542.

Sottrup-Jensen, L. (1994). Role of internal thiol esters in the α-macroglobulin-proteinase binding mechanism. Ann. N. Y. Acad. Sci. *737*, 172–187.

Sottrup-Jensen, L., Stepanik, T.M., Kristensen, T., Lonblad, P.B., Jones, C.M., Wierzbicki, D.M., Magnusson, S., Domdey, H., Wetsel, R.A., Lundwall, A. et al. (1985). Common evolutionary origin of α_2-macroglobulin and complement components C3 and C4. Proc. Natl. Acad. Sci. USA *82*, 9–13.

Sottrup-Jensen, L., Sand, O., Kristensen, L., and Fey, G.H. (1989). The α-macroglobulin bait region. Sequence diversity and localization of cleavage sites for proteinases in five mammalian α-macroglobulins. J. Biol. Chem. *264*, 15781–15789.

Starkey, P.M. and Barrett, A.J. (1973). Inhibition by α-macroglobulin and other serum proteins. Biochem. J. *131*, 823–831.

Starkey, P.M. and Barrett, A.J. (1982). Evolution of α_2-macroglobulin. The demonstration in a variety of vertebrate species of a

protein resembling human α_2-macroglobulin. Biochem. J. *205*, 91–95.

Stöcker, W., Breit, S., Sottrup-Jensen, L., and Zwilling, R. (1991). α_2-Macroglobulin from hemolymph of the freshwater crayfish *Astacus astacus*. Comp. Biochem. Physiol. B. *98*, 501–509.

Stoops, J.K., Schroeter, J.P., Bretaudiere, J.P., Olson, N.H., Baker, T.S., and Strickland, D.K. (1991). Structural studies of human α_2-macroglobulin: concordance between projected views obtained by negative-stain and cryoelectron microscopy. J. Struct. Biol. *106*, 172–178.

Strickland, D.K., Ashcom, J.D., Williams, S., Burgess, W.H., Migliorini, M., and Argraves, W.S. (1990). Sequence identity between the α_2-macroglobulin receptor and low density lipoprotein receptor-related protein suggests that this molecule is a multifunctional receptor. J. Biol. Chem. *265*, 17401–17404.

Tapon-Bretaudiere, J., Bros, A., Couture-Tosi, E., and Delain, E. (1985). Electron microscopy of the conformational changes of α_2-macroglobulin from human plasma. EMBO J. *4*, 85–89.

Thogersen, I.B., Hammes, S.R., Rubenstein, D.S., Pizzo, S.V., Valnickova, Z., and Enghild, J.J. (2002). New member of the trefoil factor family of proteins is an α-macroglobulin protease inhibitor. Biochim. Biophys. Acta *1598*, 131–139.

Travis, J. and Salvesen, G.S. (1983). Human plasma proteinase inhibitors. Annu. Rev. Biochem. *52*, 655–709.

Vogel, C.W., Finnegan, P.W., and Fritzinger, D.C. (2014). Humanized cobra venom factor: Structure, activity, and therapeutic efficacy in preclinical disease models. Mol. Immunol. *61*, 191–203.

Wang, D., Wu, K., and Feinman, R.D. (1983). Dissociability of enzyme-α_2-macroglobulin complexes. Arch. Biochem. Biophys. *222*, 117–122.

Woessner, Jr. J.F. and Nagase, H. (2000). Matrix Metalloproteinases and TIMPs. Protein Profile Series, P. Sheterline, ed. (New York, NY: Oxford University Press).

Wong, S.G. and Dessen, A. (2014). Structure of a bacterial α_2-macroglobulin reveals mimicry of eukaryotic innate immunity. Nat. Commun. *5*, 4917.

Xiao, T., DeCamp, D.L., and Sprang, S.R. (2000). Structure of a rat α_1-macroglobulin receptor-binding domain dimer. Protein Sci. *9*, 1889–1897.