Structural and functional analysis of cystatin E reveals enzymologically relevant dimer and amyloid fibril states

Received for publication, January 29, 2018, and in revised form, June 26, 2018. Published, Papers in Press, July 2, 2018, DOI 10.1074/jbc.RA118.002154

Elfriede Dall†, Julia C. Hollerweger‡, Sven O. Dahms‡, Haissi Cui§, Katharina Häussermann§, and Hans Brandstetter†1

From the †Department of Biosciences, University of Salzburg, A-5020 Salzburg, Austria and the ‡Center for Integrated Protein Science Munich, Technical University of Munich, D-85748 Munich, Germany

Edited by Wolfgang Peti

Protein activity is often regulated by altering the oligomerization state. One mechanism of multimerization involves domain swapping, wherein proteins exchange parts of their structures and thereby form long-lived dimers or multimers. Domain swapping has been specifically observed in amyloidogenic proteins, for example the cystatin superfamily of cysteine protease inhibitors. Cystatins are twin-headed inhibitors, simultaneously targeting the lysosomal cathepsins and legumain, with important roles in cancer progression and Alzheimer’s disease. Although cystatin E is the most potent legumain inhibitor identified so far, nothing is known about its propensity to oligomerize. In this study, we show that conformational destabilization of cystatin E leads to the formation of a domain-swapped dimer with increased conformational stability. This dimer was active as a legumain inhibitor by forming a trimeric complex. By contrast, the binding sites toward papain-like proteases were buried within the cystatin E dimer. We also showed that the dimers could further convert to amyloid fibrils. Unexpectedly, cystatin E amyloid fibrils contained functional protein, which inhibited both legumain and papain-like enzymes. Fibril formation was further regulated by glycosylation. We speculate that cystatin amyloid fibrils might serve as a binding platform to stabilize the pH-sensitive legumain and cathepsins in the extracellular environment, contributing to their physiological and pathological functions.

Cysteine proteases are key regulators in many physiological processes. Consequently, dysregulation of protease activity can have severe effects resulting in a variety of pathologies, including cancer and Alzheimer’s disease (1, 2). Therefore, proteases must be regulated delicately and on different levels. Cystatins are inhibitors specifically controlling the activity of cysteine proteases. The cystatin superfamily shares a common fold that is characterized by a 5-stranded antiparallel β-sheet that is enfrraping a central α-helix. It is further organized into three subfamilies: (i) stefins; (ii) cystatins; and (iii) kininogens, fetuins, and noninhibitory cystatins (3, 4). Family 1 cystatins (stefins) are mainly localized intracellularly and ubiquitously expressed in most cell types (3, 5). They are potent inhibitors of papain (family C1) and the papain-like cathepsins with $K_I$ values in the low nanomolar range (6, 7). The interaction of stefins with papain is mediated by a tripartite wedge-shaped structure formed by the N terminus (Ser1–Val39, cystatin C numbering) and two hairpin loops (loops L1 and L2). Essentially, the N terminus binds to the nonprimed side, whereas the two adjacent hairpin loops occupy the primed substrate–binding sites.

Family 2 cystatins resemble the largest subfamily of the cystatin fold, with seven members identified so far. In contrast to the stefins, selected family 2 cystatins (C, E/M, and F) harbor, in addition to their papain-binding site, a legumain binding site (8–10). Human legumain is a caspase-like cysteine protease (family C13) that mainly localizes to the endo-lysosomal system, where it plays an important function for the processing of antigens for presentation on the MHCII complex (11). On a pathophysiological level, legumain has been implicated in various disorders, including cancers and Alzheimer’s disease (12–14). Under these conditions, legumain was found translocated to the nucleus, to the cytoplasm, and extracellularly. Because of its strict specificity for cleaving after asparagine residues, it is synonymously referred to as the asparaginyl-endopeptidase (AEP) (15, 16). This strict preference is exploited by the legumain-inhibitory cystatins C, E, and F, which use a conserved Asn39 residue, localized on a reactive center loop different from the papain-inhibitory site to specifically bind to the legumain active site (9, 17). Furthermore, the interaction with legumain involves an additional legumain exosite loop (LEL) inserted between cystatin strands β3 and β4. Complex formation leads to conformational stabilization of the pH-sensitive legumain at near neutral pH. Unlike family 1 cystatins, legumain-inhibitory cystatins are secreted outside the cell and are in some cases glycosylated (10, 18–20). Whereas cystatin C is ubiquitously expressed in different human tissues, cystatin E/M is mainly

2 The abbreviations used are: AEP, asparaginyl-endopeptidase; hCE, human cystatin E; hCC, human cystatin C; RCL, reactive center loop; LEL, legumain exosite loop; SEC, size exclusion chromatography; cmk, chloromethylketone; Z, benzoyloxy carbonyl; AMC, 7-amino-4-methylcoumarin; glycosylated cystatin E; ChEs, 2-cyclohexylynamino)ethanesulfonic acid; RMSD, root mean square deviation; ThT, thioflavin T; LEXSY, Leishmania tarentolae expression system; TEM, transmission electron microscopy.

This work was supported by the Austrian Science Fund (FWF project W_01213). The authors declare that they have no conflicts of interest with the contents of this article.

© 2018 Dall et al. Published by The American Society for Biochemistry and Molecular Biology, Inc.
Cystatin E transforms to multimers with distinct functions

localized to skin epithelia, emphasizing its role in cutaneous biology (5, 10, 21). Co-localization of human cystatin E (hCE) and legumain has been reported in hair follicles (22).

Cystatins not only encode a high intrinsic variability because of their function as dual protease inhibitors but also because of their ability to transform to distinct oligomerization states upon conformational destabilization. Factors triggering this oligomerization include N-terminal truncation by proteolytic enzymes, acidic pH, heating, and point mutations. These cause dimer formation via a domain-swapping mechanism (23–25). Essentially, the N-terminal segment, comprising β1, α, and β2 up to the L1 loop, of one monomer exchanges with that of a second monomer (26). Consequently, the papain-inhibitory site becomes inaccessible, whereas the legumain-inhibitory site remains intact. Cystatin C oligomerization leads to the formation of amyloid deposits in the brain at advanced age (25). A naturally occurring L68Q variant was identified in the cerebral fluid of patients suffering from hereditary cystatin C angiopathy (Iceland disease), which accelerates this process significantly (6, 27). Similarly, N-terminally truncated cystatin C, lacking the first 10 amino acids of the native sequence, was isolated from cystatin C amyloid deposits (28). This truncation was associated with proteolytic processing by proteases released to the cerebrospinal fluid and similarly results in accelerated formation of amyloid deposits (29). Stefin B was also reported to form amyloid fibrils and is an Aβ-binding protein and therefore supposed to play a role in Alzheimer’s disease (30–32).

Both legumain and cystatins became attractive drug targets due to their relevance in different types of cancer and dementia. Among the cystatins, the family 2 cystatins became especially interesting, because of their function as dual protease inhibitors and because they are secreted to the extracellular space, where legumain and cathepsins are similarly observed under pathophysiologic conditions. Cystatin E is the most potent physiologic legumain inhibitor, binding 100-fold more tightly as compared with cystatin C (7). Thereby, it is associated with a tumor suppressor function in prostate cancer, melanoma, and oral carcinoma cells (33–35). Furthermore, cystatin E has been observed co-localized with legumain in the extracellular environment under normal and under pathophysiologic conditions (22, 36). Notably, not only co-localization but also co-trafficking of legumain together with cystatin E from outside the cell to inside a cell has been reported (37).

However, cystatin E’s physico-chemical properties, including its propensity to form dimers and higher oligomers, have hardly been studied so far. Therefore, we set out to challenge the conformation of cystatin E by different triggering factors, including N-terminal truncation, pH, and heating. Furthermore, we analyzed the effect of these destabilizing factors on its three-dimensional structure and inhibitory function.

Results

Cystatin E forms a dimer under destabilizing conditions

To mimic the effect of N-terminal proteolytic processing, we recombinantly produced an N-terminally truncated human cystatin E (ΔhCE) variant, lacking the first 10 amino acids following the mature N terminus (Δ[Arg4–Leu13]hCE; ΔhCE in Fig. S1). Size-exclusion experiments (SEC) revealed that whereas full-length WT hCE migrated at the expected elution volume of the monomeric 15-kDa protein, the N-terminally truncated ΔhCE variant showed an additional peak eluting at the size of a dimer (Fig. 1A). This observation was already the first indication for a similar tendency of hCE to oligomerize as described for other family members. As this oligomerization tendency correlated with conformational destabilizations, we next investigated the effect of heating on full-length hCE. Interestingly, we found that cystatin E monomer efficiently converted to a dimeric form upon incubation at ≈70 °C (Fig. 1A and B). Using SEC, we could determine the transition temperature where 50% conversion was reached to be 65 °C (Fig. 1B). This was further cross-confirmed by ThermoFluor experiments (64), which revealed a melting temperature of monomeric hCE of 65 °C (Fig. 1C, black curve). In this particular case, thermally induced increase in fluorescence at 65 °C corresponded to the transition to dimeric hCE rather than to protein unfolding. To highlight the relevance of these experiments, we used cystatin C as a standard to compare the thermal energy barriers for dimer formation between the two family members. Significantly, cystatin C showed qualitatively and quantitatively the same behavior as hCE with 50% conversion to dimeric human cystatin C (hCC) at 65 °C and 100% conversion at 70 °C (Fig. 1B). Similarly, we tested the effect of pH on hCE dimerization. Indeed, incubation of monomeric hCE at pH 3.0 could significantly reduce the transition temperature. SEC experiments revealed efficient conversion to dimeric hCE already at 60 °C (Fig. 1D). Additionally, using the ThermoFluor method, we observed a reduction of the transition temperature by 8 °C upon incubation at pH 3.5 (Fig. 1C, gray curve).

The cystatin E dimer is the thermodynamically preferred conformation

Based on these experiments, we concluded that conversion of monomeric to dimeric hCE is associated with an energy barrier that needs to be overcome. The energy barrier can be reduced by destabilizing conditions like N-terminal truncation, low pH, and heating. Similarly, long-term incubation at ambient temperature will have the same effect. To test whether the hCE dimer is a stable folding state by itself, we examined melting curves of monomeric and dimeric hCE (Fig. 1C). Dimeric hCE was generated by incubation of the monomer at 80 °C. Interestingly, the dimer showed an increased melting temperature, indicating that once the energy burden of dimerization has been overcome, the dimer is the structurally more stable conformation. Additionally, we found the monomer–dimer transition to be irreversible (i.e. dimeric hCE could not be converted back to monomeric hCE). Monomeric hCE thus represents a metastable state that is kinetically preferred in the folding process over the more stable dimeric hCE. Together, these observations were consistent with a cystatin C–like mode of dimerization, which is mediated by domain swapping (26).

The cystatin E dimer forms via domain swapping

To unveil the molecular mechanism of hCE dimerization, we determined the crystal structure of dimeric hCE, which was
Cystatin E transforms to multimers with distinct functions

Figure 1. Dimerization of hCE is triggered by destabilization. A, dimer formation of N-terminally truncated (ΔhCE) and WT (hCEwt) hCE was assayed following incubation at the indicated temperatures using SEC. B, hCE and hCC were incubated at the indicated temperatures before injection onto the SEC column. At 70 °C, >90% conversion to the dimeric form could be observed. C, thermal denaturation curves were collected for monomeric (hCEm) and dimeric (hCEd) hCE at the indicated pH values. Melting Temperatures (TM) could be determined to be 65 °C (transition 1) and 87 °C (transition 2) at pH 5.5 for hCEm and hCEd, respectively. Acidic pH led to a reduction of TM1 and TM2 region from one molecule into the sec-
ond. Therefore, the cystatin E monomer and dimer are two structurally distinct subspecies; the dimer is not the sole combination of two monomers.

From a structural perspective, the hCE dimer was assembled from two hCE monomer conformers that are connected via their L1 loops. When we compared the structural similarity of the two half-domains hCEE’ and hCE’E within the dimer, we observed a relatively low root mean square deviation (RMSD) of the Cα positions (0.33 Å). The deviations between the two subdomains may relate to different packing environments within the crystal lattice. Interestingly, superposition of dimeric hCE with monomeric hCE revealed an average RMSD of 0.66 Å, indicating significant structural differences in hCE monomer as compared with the corresponding dimeric subdomain (Fig. 2A). Structural differences accumulated from a number of small rearrangements that accompanied the integration of the N-terminal β1-α1-β2 region from one molecule into the sec-
ond. Therefore, the cystatin E monomer and dimer are two structurally distinct subspecies: the dimer is not the sole combination of two monomers.

When we superposed hCE and hCC monomer structures, we found an RMSD of the Cα positions of 1.3 Å (Fig. S2C). Interestingly, the difference between the hCE and hCC dimer structures is increased by 2-fold (RMSD = 2.8; Fig. S2D). Most of this bonded contacts is also reflected in a very low-solvation free energy gain ΔG upon complex formation of −43.5 kcal/M. Furthermore, the free energy of complex dissociation, ΔG^diss, of 55.4 kcal/M indicates that the dimer is thermodynamically stable.

The cystatin E dimer adopts a structurally distinct conformation

From a structural perspective, the hCE dimer was assembled from two hCE monomer conformers that are connected via their L1 loops. When we compared the structural similarity of the two half-domains hCEE’ and hCE’E within the dimer, we observed a relatively low root mean square deviation (RMSD) of the Cα positions (0.33 Å). The deviations between the two subdomains may relate to different packing environments within the crystal lattice. Interestingly, superposition of dimeric hCE with monomeric hCE revealed an average RMSD of 0.66 Å, indicating significant structural differences in hCE monomer as compared with the corresponding dimeric subdomain (Fig. 2A). Structural differences accumulated from a number of small rearrangements that accompanied the integration of the N-terminal β1-α1-β2 region from one molecule into the secondary. Therefore, the cystatin E monomer and dimer are two structurally distinct subspecies; the dimer is not the sole combination of two monomers.

When we superposed hCE and hCC monomer structures, we found a RMSD of the Cα positions of 1.3 Å (Fig. S2C). Interestingly, the difference between the hCE and hCC dimer structures is increased by 2-fold (RMSD = 2.8; Fig. S2D). Most of this
high difference can be accounted for by the flexibility of the connecting hinge region (L1–L1’). Small structural changes within this region translate into large motions in the peripheral regions (e.g. the legumain-binding site). Given the flexibility of the connecting hinge, differences in crystal packing may also contribute to differences in the relative spatial orientation of the two subdomains, both in the hCE and hCC dimer structures (25). Therefore, we analyzed the crystal packing of the hCE dimer crystals. Remarkably, we discovered the assembly of two hCE dimers to a tetramer (Fig. S2E). A similar assembly was described previously for stefin B (PDB code 2OCT) and cystatin C (PDB code 1G96). The stefin B tetramer forms via a handshake mechanism of the L2 loops that is triggered by a proline switch (Pro105 cystatin C numbering; Fig. S2). Pro105 is conserved throughout the whole cystatin family. However, the L2 loop is shortened by 3 amino acids in the family 2 cystatins (Fig. S1). Hence, a direct handshake between two L2 loops is not possible. Nevertheless, the overall assembly (nearly parallel β2-β3 strands) is similar between stefin B, cystatin C, and cystatin E and therefore suggests a conserved mechanism of oligomerization. The dimers interact via hydrophobic contacts involving residues of the former L1 loops. The hCE tetramer is visible in the crystal structure but not in SEC experiments, indicating a low affinity of this assembly.

**The cystatin E dimer is a functional legumain inhibitor**

The hCE dimer is a unique folding state that encodes two legumain-binding sites on one molecule. In contrast to monomeric hCE, the legumain-binding sites of the dimer are built up by contributions of both hCE chains (e.g. legumain-binding site 1 is built up from the RCL’ (reactive center loop) of one chain, hCE’, and the LEL of the second chain, hCE, and vice versa (Figs. 2B and 3A). Despite this large structural rearrangement, the local conformations of the legumain RCL and LEL were virtually identical in monomeric and dimeric hCE structures (Fig. 3B). Thus, we hypothesized that dimeric hCE is still functional as a legumain inhibitor. Indeed, when we incubated legumain with dimeric hCE, we observed a complete loss of its Ala-Ala-Asn-AMC substrate turnover (Fig. 3C). The crystal structure of dimeric hCE uncovered two symmetric and functional legumain-binding sites, either of which is accessible for legumain binding. To test whether a simultaneous binding of two legumain molecules by one hCE dimer (i.e. a tetrameric assembly) is sterically feasible, we prepared a model of such a complex based on the crystal structures of the hCE dimer and the legumain–hCE complex (PDB entry 4N6O (17)). First, we superposed the hCE dimer onto the legumain–hCE complex (Fig. 3D). We found virtually identical intermolecular contacts for either of the two possible docking modes (i.e. docking legumain to the hCE’ or to the hCE’E submolecule was equally well accessible). However, from these docking models, it became immediately clear that simultaneous binding to both sites is sterically impossible. When legumain binds to the hCE’ site, the free hCE’E site is in close proximity to the legumain insertion loop on legumain and vice versa. Hence, a trimer composed of one hCE dimer and one legumain (AEP) seemed to be the most likely assembly (Fig. 3D). To test this hypothesis, we performed size-exclusion experiments of a preformed legumain–hCE dimer complex. Indeed, we observed a peak at the size of a trimer, but not at the expected size of a tetramer (Fig. 3E). Additionally, SDS-PAGE revealed a 2-fold higher hCE concentration in the fractions of the legumain–hCE dimer complex (AEP–hCE) as compared with legumain–hCE monomer complex (AEP–hCEm, Fig. 3F). The legumain concentration was identical in both experiments. The difference in hCE concentration within the fractions containing the hCE–legumain complex indicates a 2:1 stoichiometry rather than 2:2. Thus, cystatin E virtually loses 50% of its legumain binding sites via domain swapping.
Cystatin E transforms to multimers with distinct functions

...demonstrates the accessible and functional legumain-inhibitory site on dimeric hCE. First, we could previously show that interaction of the LEL on cystatin E with the legumain-primed substrate-binding site has a positive effect on its thermal stability at near neutral pH, where the isolated enzyme by itself is unstable (17, 41). The same is true for the legumain–hCE dimer complex (Fig. 3G). Second, we observed pH-dependent cleavage and religation of monomeric hCE at the conserved Asn39 residue located on the legumain reactive center...
Cystatin E transforms to multimers with distinct functions

Figure 4. Glycosylation is compatible with hCE dimerization. A, top view of the hCE dimer structure in cartoon representation. The Asn112,12 residues located on the L2 loops that are prone to glycosylation are shown in sticks. B, glycosylated hCE produced in LEXSY was incubated at 20 °C (black curve) and 85 °C (dashed, black curve) for 10 min. Subsequently, both samples were injected onto an S75 10/300 GL column. Incubation at 85 °C led to a shift in the retention volume that corresponds to dimeric hCE. C, inhibition of papain and legumain by monomeric (gray) and dimeric (white) glycosylated cystatin E was assayed using fluorogenic FR-AMC (papain) and Z-AAN-AMC (legumain) substrates. Activities were normalized to control reactions harboring enzyme only (black). mAU, milli-absorbance units.

loop of hCE by legumain (17). Cleavage happens at pH <5.5, and religation is observed upon incubation of cleaved hCE at pH >6.5. We observed the same pH-dependent behavior when dimeric hCE was used (Fig. 3H). Together, these observations are in nice agreement with the crystal structure of dimeric hCE that uncovered structurally and functionally uncompromised legumain-binding sites.

The cystatin E dimer is not functional as papain inhibitor

In contrast to the legumain-inhibitory site, the papain-inhibitory site structurally differs in monomeric and dimeric hCE (Fig. 2, A and B). Upon dimerization, the L1 loop, which is essential for binding to the papain S1’ site, was incorporated into the newly formed βII-βIII sheet. Thus, loop L1 is not available for binding anymore, and the papain-inhibitory site is destroyed by the dimer interface. In agreement with this observation, dimeric hCE did not show inhibition of papain and human cathepsin S in a peptide substrate hydrolysis assay (Fig. 3C). Consequently, dimerization of cystatin E led to 100% loss of its cathepsin-inhibitory function.

Glycosylation on the L2 loop is compatible with dimerization

Native cystatin E harbors an N-glycosylation site on its L2 loop (Asn112). Because both glycosylated and nonglycosylated hCE were previously reported in vivo, we were interested in the relevance of glycosylation for hCE dimerization (20, 42). The crystal structure of the hCE dimer suggested no negative effect of glycosylation on dimer formation (Fig. 4A). To test whether this is the case, we incubated glycosylated hCE produced in the Leishmania tarentolae expression system (LEXSY) at 85 °C and subjected it to SEC. Indeed, we observed a shift corresponding to dimeric, glycosylated hCE (Fig. 4B). In agreement with the domain-swapping mechanism, this dimer similarly inhibited legumain but not papain in a fluorescent peptide substrate assays (Fig. 4C). Furthermore, ThermoFluor experiments of glycosylated hCE revealed a similar thermal unfolding behavior as observed for unglycosylated hCE (Fig. S3).

Cystatin E is resistant to heterodimerization

Because the overall structures of stefin B and cystatin C and E dimers are superficially similar, we questioned whether dimerization might not be restricted to homodimer formation but rather open to the possibility of heterodimerization as well. To test this, we performed pulldown assays using a cystatin E construct carrying a C-terminal Strep-tag and WT cystatin C. Both proteins were co-incubated at 70 °C to ensure efficient domain swapping. Subsequently, we immobilized hCE on a Strep-Tactin resin via its Strep-tag. If heterodimerization had occurred, we would have expected co-purification of the untagged hCC. However, using SDS-PAGE, we could not detect co-migration of hCC; solely hCE homodimer had bound (Fig. S4). This observation can probably be understood by the low sequence identity of only 30% between cystatin E and C. Sequence variations on the α1 helix, which is a critical connecting element, may result in steric clashes, thereby preventing heterodimerization. Inspection of the hCE and hCC monomer structures revealed, for example, a hCE-Ala29PhehCC variation that may be in steric conflict with Ser114-Gln-Leu116 (85) on hCE. Thus, the affinity for homodimer formation is supposedly much higher than for heterodimerization. It might be possible to outcompete this affinity difference by having a large excess of hCC relative to hCE.

Cystatin E forms amyloid fibrils

Cystatin E domain swapping results in the formation of an energetically more favorable conformation. In the simplest situation, two domain-swapped monomers assemble to form a dimer. However, domain swapping is not restricted to dimerization, but in principle also allows the formation of multimers via concerted swapping reactions (chain reaction). The existence of such multimers was previously reported for stefin B and cystatin C and resulted in the assembly of highly ordered fibrils (26, 27, 30). Thermal denaturation curves of monomeric cystatin E revealed a first transition/partial unfolding between 60 and 70 °C, which corresponded to dimer formation (Fig. 1C). Additionally, we observed a second transition between 80 and 90 °C for both monomeric and dimeric hCE, pointing toward the presence of another conformational state like fibrils, analogous to cystatin C. To test this hypothesis, we incubated cystatin E at a temperature above the second transition point (90 °C). Thereby, we observed the formation of an insoluble protein pellet. To test whether this pellet contains misfolded, aggre-
Cystatin E transforms to multimers with distinct functions

Gated protein or rather structured amyloid fibrils, we performed X-ray diffraction experiments. Interestingly, we observed two diffraction maxima at 10 and 4.7 Å resolution, which are indicative of cross-β structures, which typically build up amyloid fibrils (Fig. 5A). Furthermore, we could confirm the presence of fibrils using thioflavin T, a dye specifically binding to amyloid structures (Fig. 5B). Even more, negative staining TEM experiments also revealed the presence of fibril structures (Fig. S5). In agreement with thermal denaturation curves, fibril formation is more efficient at 90 °C as compared with 80 °C because the energy barrier is only partially overcome at 80 °C (Figs. 1C and 5B).

hCE fibrils contain functional protein

Based on the assumption that fibril formation is a result of concerted domain-swapping reactions, it seemed possible that hCE fibrils contained structured, functional protein. Because domain swapping did not compromise the functionality of the legumain-binding site, we tested whether hCE fibrils can, like dimeric hCE, inhibit legumain activity. Indeed, upon co-incubation of legumain with hCE fibrils, we observed complete inhibition of the Ala-Ala-Asn-AMC substrate turnover (Fig. 6A). Specifically, we co-incubated monomeric and dimeric hCE, inhibit legumain activity. Indeed, upon co-incubation of legumain with hCE fibrils, we observed complete inhibition of the Ala-Ala-Asn-AMC substrate turnover (Fig. 6A). Furthermore, binding of ThT was measured as an increase in fluorescence detected at 482 nm. a.u., arbitrary units. Error bars, S.D.

hCE fibrils inhibit papain(-like) enzymes

In principle, both monomeric and dimeric hCE are suitable building blocks for legumain-inhibitory hCE fibrils. However, the situation is different for papain. Only monomeric hCE is a functional papain inhibitor. Surprisingly, co-incubation of papain with hCE fibrils led to a complete inhibition of its enzymatic activity (Fig. 6A). Thus, hCE fibrils must contain, to some extent, monomeric hCE protein, which has not undergone domain swapping. Presumably, hCE fibrils are built up by domain-swapped monomers but are at the same time heterogeneously decorated with monomeric WT protein (Fig. 10).

Dimeric hCE further converts to fibrils

Thermal denaturation experiments uncovered that dimeric hCE has a higher thermal stability as compared with monomeric hCE. However, both species showed a transition at T >80 °C, indicating that dimeric hCE can similarly be converted to amyloid fibrils when the energy burden is significantly lowered (e.g. by further increase in temperature) (Fig. 1C). Indeed, we found fibril formation of dimeric hCE upon incubation at 90 °C, as evidenced in a thioflavin T (ThT) test (Fig. 5B). Therefore, the dimer is not a dead end, but most likely a folding intermediate on the route to multimers. Because both monomeric and dimeric hCE are capable of fibril formation, both might potentially serve as a building block for fibrils. However, based on ThermoFluor experiments, we would rather suggest that the dimer is the critical intermediate. To study this hypothesis, we set up a fibril nucleation assay where we tested the ability of monomeric and dimeric hCE to extend preformed fibrils. Specifically, we co-incubated monomeric and dimeric hCE with purified fibrils at 90 °C for 10 min. Interestingly, we observed an increase in ThT signal for dimeric hCE spiked with fibrils as compared with the control reaction without preaddition of fibrils (Fig. 7). Thus, monomerization of dimeric hCE can be triggered by providing preformed fibrils as a “folding” template.

Fibril formation is pH-dependent and incompatible with glycosylation

Domain swapping is triggered by conformational destabilization of monomeric hCE. Besides heating, we could identify low pH as a trigger factor for hCE dimerization. Similarly, we tested the effect of acidic pH on fibril formation and thereby found an about 4-fold increase in ThT signal following incubation at pH 3.0 (Fig. 8A). Furthermore, ThermoFluor experiments revealed a reduction of the fibril transition temperature by ~8 °C at pH 3.5 (Fig. 1C). Consequently, changing the pH environment will have an effect on the hCE oligomerization state. Along that line, we were also curious about the effect of

Figure 5. Cystatin E forms cross-β structures. A, X-ray diffraction experiments of insoluble hCE pellet revealed two rings at 10 and 4.7 Å resolution, which are characteristic for cross-β structures. B, monomeric (light gray) and dimeric (dark gray) hCE were incubated at 20, 80, and 90 °C for 10 min. Subsequently, binding of ThT was measured as an increase in fluorescence detected at 482 nm. a.u., arbitrary units. Error bars, S.D.
glycosylation on hCE multimerization. Although dimerization is possible for the glycosylated variant, multimerization might be negatively affected because of potential steric conflicts. Indeed, we did not observe fibril formation of glycosylated cystatin E (glyco-hCE) upon incubation at elevated temperatures (Fig. 8, B and C). This suggested to us a regulatory function of glycosylation for hCE function.

**Discussion**

Cystatin E is an intrinsically stable protein with a melting temperature >60 °C. However, destabilization does not lead to hCE denaturation as would be the case for most other proteins, but results in the transition to an energetically more favorable conformation (Fig. 9). This second folding state, the hCE dimer, represents a structurally distinct conformation with increased conformational stability. The energy barrier that needs to be overcome to allow for domain swapping can be reduced by triggering factors. Within this study, we could identify N-terminal truncation, acidification, and heating as accelerators of dimerization (Fig. 9). Similarly, proteolytic processing has previously been reported for hCC, presumably catalyzed by leukocyte elastase, and led to faster dimerization and amyloid deposit formation (25). Along that line, a different but similar strategy of controlling the oligomerization state via proteolytic processing evolved in another member of the family 2 cystatins, cystatin F. Cystatin F harbors two additional cysteine residues, which form two intermolecular disulfide bonds and thereby connect two cystatin F monomers. That way, the papain-inhibitory site is blocked in an inhibitory latent dimeric cystatin F. However, the legumain-inhibitory site is freely accessible. Activation of cystatin F is mediated by disulfide reduction or proteolysis in the N-terminal region (19, 43). Although nothing is known so far about proteolytic processing of hCE in vivo, a similar event seems likely because of the flexible loop structure of the N-terminal region, which makes it prone to proteolytic processing. In vitro, we could previously demonstrate processing of hCC and hCE on their legumain RCL by legumain (17). Similarly, destabilization on this site may have a negative effect on the stability of monomeric and dimeric hCE. Interestingly, this cleavage is most efficient at acidic pH, where dimerization is also more easily achievable. Acidification is a common phenomenon in the brain of patients suffering from dementia, emphasizing its relevance for domain swapping and consequent aggregation (44, 45). Likewise, the L68Q hCC variant found in the cerebral fluid of patients suffering from Iceland disease has a higher tendency to form dimers and multimers. Replacement of the hydrophobic Leu⁶⁸ by the hydrophilic glutamine causes local destabilization of the monomer, thereby also reducing the energy barrier for domain swapping (25). Such a destabilizing mutation has not
yet been identified for hCE; however, comparable natural variations may exist. Domain swapping results in a complete loss of hCE’s papain-inhibitory activity and a 50% loss of the legumain binding capacity. Consequently, domain swapping of cystatins in general will lead to an increase in protease activity due to loss of the inhibitors. Domain swapping thus provides another strategy to regulate proteolytic activity and can cause dysregulation of protease activity under pathological conditions. Legumain can stabilize cystatin E in its monomeric and dimeric state by binding to the actual conformation and thereby preventing further domain swapping. However, papain(-like enzymes) can only bind and retain the cystatin E monomer.

Although the conversion of monomeric to dimeric hCE is kinetically trapped, dimeric hCE can further convert to fibrils. Thus, domain swapping seems to be key for multimer formation. Whereas the addition of preformed fibrils to dimeric hCE led to an increase in ThT fluorescence, monomeric hCE was unaffected. This further confirms the relevance of the hCE dimer as a folding intermediate on the route to higher cystatin multimers. Additionally, it implies some prion-like behavior encoded in hCE. Because the hCE dimer has already been prone to domain swapping, it seems plausible that it is a better acceptor for a folding template than the monomer. However, it remains unclear whether monomeric hCE directly converts to fibrils if enough energy is supplied (1-step mechanism) or whether the dimer is formed first but immediately converted to fibrils (2-step mechanism; Fig. 10). A number of intermediate states have been described for other cystatins already, including a molten globule state or a tetramer (46, 47). Therefore, the full picture presumably is more complex and not restricted to one intermediate.

Amyloid proteins are generally considered to lose their native conformation while forming well-ordered cross-β structures (48). This allows them to assemble into large structures containing many copies of the same molecule. A prominent example of such a protein is the Aβ-peptide, which plays a central role in the pathology of Alzheimer’s disease (49). Surprisingly, we could show that hCE amyloid fibrils contain biologically functional protein. Whereas the ability of hCE fibrils to inhibit legumain is in agreement with both domain-swapped and monomeric hCE, papain inhibition is only possible with monomeric hCE. Hence, we concluded that hCE amyloid fibrils contain domain-swapped hCE as a building block to fibril formation; however, they are heterogeneously decorated with monomeric hCE (Fig. 10). Domain swapping may not be 100% efficient, explaining how some monomeric protein remains to be integrated into fibrils. Survival of monomeric protein is presumably time-dependent. Consequently, longer incubation may result in less monomeric protein and less papain inhibition. Based on these observations, we assume that hCC fibrils also contain, to some extent, functional protease inhibitor.

Cystatin fibrils might serve as a binding platform to stabilize the pH-sensitive legumain and cathepsins in the extracellular environment, thereby increasing their lifetime. Because cystatins are reversible inhibitors, fibrils might provide a strategy of storing enzymes for later action. Although we did not observe heterodimerization between family 2 members, the
formation of mixed fibrils seems possible. Protodimers of hCE and hCC may assemble into mixed larger structures.

Unlike the family 1 cystatins, family 2 cystatins are frequently glycosylated. Human cystatin E harbors an N-glycosylation site on the L2 loop, which is compatible with dimerization but abolished fibril formation. From a structural perspective, modification on the L2 loop is in principle compatible with domain swapping but in steric conflict with the formation of large fibrils, where proteins need to be packed together tightly. Interestingly, both glycosylated and unglycosylated hCE were observed in vivo (e.g. in breast cancer cell lines or cystatin E–overexpressing HEK293 cells), suggesting different (pathophysiological) functions for both variants (10, 20, 37). Whereas glycosylated hCE will only be present as monomer or dimer, unglycosylated hCE can potentially also be converted to amyloid fibrils. Moreover, human cystatin C also harbors a predicted glycosylation site, but different from the one present in hCE. Cystatin C harbors an O-glycosylation site on its N-terminal region (Ser2). Glycosylation at this site will probably have an effect on dimerization, because this part of the cystatin molecule is directly involved in domain swapping. Indeed, there is indirect evidence from the literature that N-terminal glycosylation is incompatible with or at least restricting hCC dimerization (50). Whereas cystatin C may lose the O-glycosylation site by N-terminal proteolytic cleavage, the hCE N-glycosylation is not affected because of its location on the L2 loop (28, 51). However, (de)glycosidases might play a role as regulatory enzymes, switching glycosylated hCE to unglycosylated hCE. N-terminally truncated cystatin C was isolated from cystatin C amyloid deposits and results in accelerated formation of amyloid deposits (29).

Recently, cystatin E was also identified in the cerebrum (52). Thus, it is attractive to speculate about a potential role of domain-swapped cystatin E in neuronal function.

**Experimental procedures**

**Preparation of proteins**

WT hCE and hCC constructs lacking the N-terminal signal sequence were cloned into the pET22b(+) vector (Novagen) as described earlier (17). A truncated version of hCE (ΔhCE) was prepared using a forward primer carrying an Ncol restriction site (5′-ATGCCTAGGATCCGCGACGACCCGACGGTG-3′) and a reverse primer carrying an XhoI restriction site (5′-ACGTCTCGAGCATCTGCACACAGGTT-3′). The construct has residues Arg4–Leu13 deleted and, due to the employed cloning strategy, starts with the double mutation R14M/D15E. Full-length WT hCE, N-terminally truncated hCE, and hCC had the native signal peptide replaced by the pelB leader sequence present on the pet22b(+) vector to allow expression to the *Escherichia coli* periplasmic space. During secretion, the signal peptide was removed, thereby liberating the new Arg^4^ N terminus of full-length WT hCE (hCC numbering), Met^14^ of the truncated ΔhCE variant, and Ser^1^ of hCC. Additionally, the expression constructs carried a C-terminal His$_6^\text{c}$ tag for purification. Furthermore, another WT cystatin E construct carrying a C-terminal His$_6^\text{c}$ tag followed by a Strep-tag was prepared. For that purpose, a WT cystatin E construct was used as a template, and the Strep-tag was introduced following
a protocol based on the inverse-PCR method. As primers, we used TCAGTTCGAAGTGAAGATTCCGGTCTGAACAA-
AGCCCGAAAG (forward) and GGCTGACCGTTCTG-
GGTGTTGCTGTCGACATTGCC (reverse). Correct-
ness of expression constructs was confirmed via DNA sequencing by Eurofins MWG Operon (Martinsried, Germany).

Cystatin E and C constructs were expressed in E. coli Bl21(DE3) cells. Briefly, the expression plasmid was trans-
formed into Bl21(DE3) cells. For large-scale expression, cells were grown in 2-liter flasks filled with 600 ml of lysogeny broth medium (Carl Roth, Karlsruhe, Germany) supplemented with 100 μg/ml ampicillin at 37 °C with agitation at 220 rpm until an
A$_{600}$ of 0.8–1.0 was reached. Expression was induced at 25 °C by the addition of 1 μl isopropyl β-D-1-thiogalactopyranoside. After overnight expression, cells were harvested by centrifuga-
tion (10 min, 4000 rpm, 4 °C) and frozen. The periplasmic frac-
tion containing recombinant proteins was extracted by cold osmotic shock. Briefly, frozen cell pellets of 1.8 liters of expres-
sion were gently resuspended in 150 ml of lysis buffer
composed of 20 mM Tris, pH 7.5, 300 mM NaCl, and 10 mM
nitrilotriacetic acid Superflow resin (Qiagen, Hilden, Germany).

After overnight expression, cells were harvested by centrifuga-
tion (10 min, 4000 rpm, 4 °C) and frozen. The periplasmic frac-
tion containing recombinant proteins was extracted by cold osmotic shock. Briefly, frozen cell pellets of 1.8 liters of expres-
sion were gently resuspended in 150 ml of lysis buffer
composed of 20 mM Tris, pH 7.5, 300 mM NaCl, and 10 mM
nitrilotriacetic acid Superflow resin (Qiagen, Hilden, Germany).

To analyze the oligomerization state of different hCE vari-
ants after different treatments, 250 μl of sample was loaded on an
S75 10/300 GL column (GE Healthcare) which was
activated as
buffer (50 mM citric acid, pH 5.5, 100 mM NaCl) containing 100
molar ratio of 1:1 (1 AEP and 1 hCE dimer) at pH 5.5 for 10
min on ice before injection of the sample on a S200 10/300 GL
column. The hCE dimer was prepared by incubation of mono-
meric hCE at 80 °C for 10 min. Similarly, legumain was incu-
bated with monomeric hCE at a molar ratio of 1:2 (1 AEP and 2
hCE monomers) for 10 min at pH 5.5. To calculate molar con-
centrations of monomeric and dimeric hCE, a molecular mass of
15 and 30 kDa, respectively, was assumed. Both samples con-
tained equal amounts of cystatin E either in monomeric or
dimeric state. For all samples investigated, fractions were col-
clected and analyzed on SDS-PAGE.

**ThermoFluor assays**

Thermal denaturation curves of different protein variants
after different treatments were determined using the Thermo-
Fluor method (54). Briefly, 1 mg/ml protein sample containing
50 × SYPRO Orange (Invitrogen) was added in a 1:10 ratio to
22.5 μl of assay buffer composed of 50 mM citric acid pH 5.5/3.0
and 100 mM NaCl. Thermal denaturation curves were collected
in a 7500 real-time PCR system (Applied Biosystems) from 20
to 95 °C. The fluorescence data were analyzed as described pre-
viously (55). Protein samples investigated were monomeric
hCE and dimeric hCE, which was prepared by incubation of
monomeric hCE at 80 °C for 10 min. To test the stability of AEP
in complex with monomeric hCE and dimeric hCE, a complex
was prepared at pH 5.5 by mixing AEP and inhibitor in a 1:1
molar ratio, assuming 30 kDa as the molecular mass of dimeric
hCE. Stability of AEP only and in complex with hCE was
assayed in a buffer composed of 50 mM MES, pH 6.5, and 100
mM NaCl. To test the effect of hCE fibrils on AEP stability, 1 μl
of washed fibrils were added to the assay buffer (50 mM MES,
pH 6.5, 100 mM NaCl) before the addition of AEP premixed
with SYPRO Orange. Fibril preparation is described under “Co-
precipitation assay.”

**Inhibition assays**

Inhibition of WT legumain was tested in legumain assay
buffer (50 mM citric acid, pH 5.5, 100 mM NaCl) containing 100
μM benzylxoycarbonyl-Ala-Ala-Asn-7-amino-4-methylcou-
marin substrate (Z-AAN-AMC; Bachem). Assays were carried
out in an Infinite M200 plate reader (Tecan). Briefly, the assay
buffer was preincubated with 8 nM cystatin, followed by the
addition of 4 nM legumain. Increase in fluorescence was mea-
sured at 460 nm upon excitation at 380 nm at 37 °C. Inhibition
of papain (EC 3.4.22.2; Merck, Darmstadt, Germany) and recombinant human cathepsin S was assayed in the same assay
buffer containing 100 μM Z-FR-AMC (Bachem). The assay
buffer was preincubated with 100 nM cystatin, followed by the
addition of 4 nM legumain. Increase in fluorescence was mea-
sured at 460 nm upon excitation at 380 nm at 37 °C. Inhibition
of papain (EC 3.4.22.2; Merck, Darmstadt, Germany) and recombinant human cathepsin S was assayed in the same assay
buffer containing 100 μM Z-FR-AMC (Bachem). The assay
buffer was preincubated with 100 nM cystatin, followed by the
addition of 50 nM papain or cathepsin S, and fluorescence was
similarly recorded at 460 nm. Dimeric cystatin E was prepared
by incubation of monomeric cystatin E at 80 °C for 10 min. The
sample was filtered to remove higher oligomers. Similarly, gly-
cosylated cystatin E was also incubated at 80 °C for 10 min to
generate the dimeric form. All experiments were carried out in triplicate.

**Determination of Kᵢ values**

Inhibition constants of monomeric and dimeric hCE toward
legumain were determined in assay buffer composed of 50 mM
citric acid, pH 5.5, 100 mM NaCl, and 0.05% Tween 20 using the
Morrison equation for tight binding inhibitors (56). First, the $K_m$ value of legumain toward the Z-AAN-AMC substrate was determined in assay buffer containing 3–250 μM substrate. The reactions were started by the addition of 2 nM enzyme. The $K_m$ was calculated to be 52 μM under these assay conditions. In the next step, the assay buffer containing 100 μM Z-AAN-AMC substrate was preincubated with increasing concentrations of hCE ranging from 0.01 to 5 nM, and the reaction was started by the addition of 2 nM legumain. Fluorescence was monitored at 460 nm and 37 °C for 10 min. The velocity of substrate turnover was calculated as fluorescence units/s, and the data points were fit to the Morrison equation using GraphPad Prism version 7.0 (GraphPad Software, La Jolla, CA). To calculate molar concentrations of dimeric hCE, a molecular mass of 30 kDa was assumed. All measurements were performed in triplicates.

**Crystallization and structure solution of dimeric hCE**

Dimeric hCE was prepared by incubation of 1 mg/ml WT hCE at 80 °C for 10 min. Subsequently, the protein sample was filtered to remove higher oligomers and subjected to SEC using a S75 10/300 GL column equilibrated in a buffer composed of 20 mm citric acid, pH 5.5, and 50 mM NaCl. Fractions containing dimeric protein were concentrated to ~30 mg/ml final concentration using Amicon Ultra centrifugal filter units (molecular weight cutoff 10,000; Millipore). Initial crystallization screening was performed in a sitting-drop vapor diffusion setup. 0.4 μl of concentrated dimeric hCE were mixed with 0.4 μl of screen solution (Hampton Index HT or JBScreen Classic) and equilibrated with 60 μl of reservoir solution in 96-well Intelli Plates (Art Robbins Instruments) at 20 °C. After 1–2 weeks, crystals were observed in a condition composed of 20% PEG 8000 and 0.1 M Ches, pH 9.5. Crystals were harvested after stepwise addition of a cryo-protectant solution containing 22% PEG 8000, 0.1 M Ches, pH 9.5, and 20% glycerol and flash-frozen in liquid nitrogen. A native data set was collected at 100 K on beamline ID33-2 (ESRF, Grenoble) equipped with a Pilatus3_2 M detector to a resolution of 2.8 Å. 720 images were collected at a wavelength of 0.8726 Å at 0.5° oscillation range.

Data processing was performed utilizing iMOSFLM (57) and SCALA from the CCP4 program suite (58). An initial search model was prepared using the crystal structure of monomeric hCE (PDB entry 4N6L) by removing the N-terminal region up to loop L1. This model was used as an initial search model for molecular replacement using PHASER (59). Repeated cycles of manual rebuilding in COOT (60) and refinement using phenix.refine (61) were carried out. The atomic coordinates and experimental structure factors have been deposited in the Protein Data Bank with accession code 6FK0. PyMOL (62) was used to create figures illustrating structures.

**Molecular modeling**

A model of dimeric hCE in complex with legumain was created using Topmatch (63). Specifically, the crystal structure of the legumain–hCE complex (PDB entry 4N6O) served as a template to align the structure of dimeric hCE.

**Proteolysis and ligation assay**

To test proteolysis at the P1-Asn39 residue on cystatin E, monomeric and dimeric cystatin E were incubated with legumain in a 1:2 molar ratio (1 legumain and 2 hCE; assuming a molecular mass of 15 kDa for both monomeric and dimeric hCE) at pH 4.0 at 37 °C until 80% turnover was observed, as judged by SDS-PAGE. Subsequently, the pH was shifted to 7.5, and the samples were incubated for 1 h at 37 °C. As control samples, we used legumain only and monomeric/dimeric hCE only. Progress of proteolysis and ligation was investigated after different time points via SDS-PAGE. Dimeric hCE was prepared from N-terminally truncated hCE.

**X-ray diffraction to test amyloid fibril formation**

Monomeric hCE was incubated at 90 °C for 10 min at 20 mg/ml protein concentration. Subsequently, the insoluble fraction potentially containing amyloid fibrils was harvested by centrifugation (16,000 × g, 10 min, 4 °C). To wash off residual monomeric protein, the pellet was resuspended in double-distilled H2O and again harvested by centrifugation. The supernatant was discarded and washed another two times. The pellet was then mounted at the edge of a quartz glass capillary, and diffraction was assayed in house using a Bruker Microstar rotating anode generator mounted with a Mar345dtb detector.

**Testing pH dependence of fibril formation**

To test for formation of amyloid fibrils, cystatin E was incubated at 80 °C for 10 min at a protein concentration of 10 mg/ml in a buffer composed of 50 mM citric acid and 100 mM NaCl (pH range 3–6) or 50 mM Tris and 100 mM NaCl (pH 7.0). Afterward, samples were put on ice for at least 10 min before testing fibril formation using a ThT test. Briefly, a ThT stock solution was prepared by dissolving 8 mg of ThT (Sigma-Aldrich) in 10 ml of PBS buffer. A working solution was prepared freshly each day by diluting the ThT stock solution 1:50 in PBS buffer. 24 μl of working solution were mixed with 1 ml of sample in a 386-well black polystyrene plate (Corning), and fluorescence was measured for 2 min at 25 °C with excitation at 440 nm and emission at 482 nm. Mean and S.D. values from three measurements were calculated.

**Testing the effect of glycosylation on fibril formation**

To test the effect of L2 loop glycosylation of cystatin E on the formation of higher oligomers, we incubated cystatin E produced in LEXSY at 90 °C for 10 min at a concentration of 20 mg/ml in a buffer composed of 50 mM citric acid, pH 5.5, and 100 mM NaCl to induce oligomerization. Subsequently, the sample was incubated on ice for at least 10 min before analysis in a ThT assay.

**Testing inhibition of legumain and papain by hCE fibrils**

Cystatin E fibrils were prepared by incubation of a 10 mg/ml sample at pH 3.0 and 80 °C for 10 min. The insoluble fraction was harvested by centrifugation (16,000 × g, 10 min); resuspended in washing buffer composed of 50 mM citric acid, pH 5.5, and 100 mM NaCl; and centrifuged again. This washing step was repeated three times. Afterward, the fibrils were resus-
Cystatin E transforms to multimers with distinct functions

pended in 30 µl of wash buffer. To test inhibition of legumain and papain, the substrate solution described before was preincubated with 1 µl of fibrils, and the reaction was started by the addition of the enzyme. In each case, the final reaction volume was 50 µl. Control reactions contained only the enzyme in substrate solution. All reactions were measured in triplicate.

Co-precipitation assay

Cystatin E fibrils were prepared by incubation of monomeric hCE at a 20 mg/ml concentration at 90 °C for 10 min. The insoluble fraction was harvested by centrifugation (16,000 × g, 10 min). To remove any residual monomeric or dimeric protein contaminants, the pellet was washed three times with a buffer composed of 50 mM citric acid, pH 5.5, 100 mM NaCl. The final pellet was resuspended in 20 µl of wash buffer. 5 µl of fibrils were mixed with 5 µl of legumain (0.1 mg/ml final concentration) and incubated for 10 min at 20 °C. Afterward, the insoluble fraction was harvested by centrifugation and washed two times with washing buffer. Control reactions contained fibrils only, fibrils + AEP precomplexed with the covalent acetyl-YVAD-cmk, and legumain only.

Testing fibril formation by the addition of preformed fibrils

Cystatin E fibrils were prepared and washed as described above. Dimeric hCE was prepared by incubation of monomeric protein at 80 °C for 10 min, followed by filtering. Monomeric and dimeric cystatin E (10 µl, 10 mg/ml) were supplemented with 1 µl of cystatin E fibrils and incubated at 90 °C for 10 min. Control samples contained fibrils only or cystatin E monomer/dimer only. Afterward, samples were incubated on ice for at least 10 min before setting up a ThT assay. The reaction was set up in a buffer composed of 50 mM citric acid, pH 5.5, and 100 mM NaCl.

Cystatin E/C heterodimerization assay

To test heterodimerization of cystatins E and C, cystatin E carrying a C-terminal Strep-tag was co-incubated with cystatin C at 70 °C for 10 min. Following a further 10-min incubation on ice, the sample was loaded on a Strep-Tactin Sepharose resin (IBA GmbH, Göttingen, Germany) pre-equilibrated in wash buffer (100 mM Tris, pH 8.0, 150 mM NaCl). The flow-through was collected, and the resin was washed five times with wash buffer. Finally, bound protein was eluted by applying five times 25 mM NaOH for 40 s. TEM imaging was performed using a Philips CM 100 transmission microscope operating at 100 kV. Images were acquired using an AMT 4 × 4 megapixel CCD camera. Imaging was performed at ×28,500 magnification.

Author contributions—E. D., J. C. H., and H. B. designed the experiments. K. H. and H. C. performed the TEM experiments. S. O. D. designed the X-ray experiment to test for fibril formation and contributed with valuable discussions. E. D. and J. C. H. performed all other experiments. E. D. and H. B. wrote the paper, and all authors reviewed the manuscript.

Acknowledgments—We thank Hendrik Dietz (Center for Integrated Protein Science Munich), for providing access to the TEM facility, Michael Groll (Center for Integrated Protein Science Munich) for support, Wait Tuck Soh for providing recombinant cathepsin S, and Boris and Dusan Turk (Institute Jožef Stefan, Ljubljana) for valuable discussions.

References

1. Mason, S. D., and Joyce, J. A. (2011) Proteolytic networks in cancer. Trends Cell Biol. 21, 228–237 CrossRef Medline

2. Hasanbasic, S., Jahic, A., Karahmet, E., Sejranic, A., and Prnjavorac, B. (2016) The role of cysteine protease in Alzheimer disease. Mater. Soc. 28, 235–238 CrossRef Medline

3. Turk, V., and Bode, W. (1991) The cystatins: protein inhibitors of cysteine proteases. FEBS Lett. 285, 213–219 CrossRef Medline

4. Hedrich, I., Lottaz, D., Meyer, K., Yiallouros, I., Jahnen-Dechent, W., Stöcker, W., and Becker-Pauly, C. (2010) Fetuin-A and cystatin C are endogenous inhibitors of human meprin metalloproteases. Biochemistry 49, 8599–8607 CrossRef Medline

5. Abrahamson, M., Barrett, A. J., Salvesen, G., and Grubb, A. (1986) Isolation of six cysteine proteinase inhibitors from human urine: their physicochemical and enzyme kinetic properties and concentrations in biological fluids. J. Biol. Chem. 261, 11282–11289 Medline

6. Turk, V., Stoka, V., Vasiljeva, O., Renko, M., Sun, T., Turk, B., and Turk, D. Cysteine cathepsins: from structure, function and regulation to new frontiers. Biochim. Biophys. Acta 1824, 68–88

7. Rawlings, N. D., Barrett, A. J., and Bateman, A. (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. 40, D343–D350 CrossRef Medline

8. Cheng, T., Hitomi, K., van Vlijmen-Willems, I. M., de Jongh, G. J., Yamamoto, K., Nishi, K., Watts, C., Reinheckel, T., Schalkwijk, J., and Zeeuw, P. L. (2006) Cystatin M/E is a high affinity inhibitor of cathepsin V and cathepsin L by a reactive site that is distinct from the legumain-binding site: a novel clue for the role of cystatin M/E in epidermal cornification. J. Biol. Chem. 281, 15893–15899 CrossRef Medline

9. Alvarez-Fernandez, M., Barrett, A. J., Gerhartz, B., Dando, P. M., Ni, J., and Abrahamson, M. (1999) Inhibition of mammalian legumain by some cystatins is due to a novel second reactive site. J. Biol. Chem. 274, 19195–19203 CrossRef Medline

10. Ni, J., Abrahamson, M., Zhang, M., Fernandez, M. A., Grubb, A., Su, J., Yu, G. L., Li, Y., Parmelee, D., Xing, L., Coleman, T. A., Gentz, S., Thotakura, R., Nguyen, N., Hesselberg, M., and Gentz, R. (1997) Cystatin E is a novel...
human cysteine proteinase inhibitor with structural resemblance to family 2 cystatins. J. Biol. Chem. 272, 10853–10858 CrossRef Medline
11. Manoury, B., Hewitt, E. W., Morrice, N., Dando, P. M., Barrett, A. J., and Watts, C. (1998) An asparaginyl endopeptidase processes a microbial antigen for class II MHC presentation. Nature 396, 695–699 CrossRef Medline
12. Haugen, M. H., Johansen, H. T., Pettersen, S. J., Solberg, R., Brix, K., Flatmark, K., and Maaløe, G. M. (2013) Nuclear legumain activity in colorectal cancer. PLoS ONE 8, e52980 CrossRef Medline
13. Liu, C., Sun, C., Huang, H., Janda, K., and Edgington, T. (2003) Overexpression of legumain in tumors is significant for invasion/metastasis and a candidate enzymatic target for prodrug therapy. Cancer Res. 63, 2957–2964 Medline
14. Zhang, Z., Obianyo, O., Dall, E., Du, Y., Fu, H., Liu, X., Kang, S. S., Song, M., Yu, S. P., Cabrele, C., Schubert, M., Li, X., Wang, J. Z., Brandstetter, H., and Ye, K. (2017) Inhibition of β-secretase improves cognitive functions in mouse models of Alzheimer’s disease. Nat. Commun. 8, 14740 CrossRef Medline
15. Li, D. N., Matthews, S. P., Antoniou, A. N., Mazzeo, D., and Watts, C. (2003) Multistep autoactivation of asparaginyl endopeptidase in vitro and in vivo. J. Biol. Chem. 278, 38980–38990 CrossRef Medline
16. Vidmar, R., Vizoviček, M., Turk, D., Turk, B., and Fonovič, M. (2017) Protease cleavage site fingerprinting by label-free in-gel degradomics reveals pH-dependent specific switch of legumain. EMBO J. 36, 2455–2465 CrossRef Medline
17. Dall, E., Fegg, J. C., Briza, P., and Brandstetter, H. (2015) Structure and mechanism of an aspartimide-dependent peptidase ligase in human legumain. Angew. Chem. Int. Ed. Engl. 54, 2917–2921 CrossRef Medline
18. Taupin, P., Ray, J., Fischer, W. H., Suhr, S. T., Hakansson, K., Grubb, A., and Gage, F. H. (2000) FGF-2-responsive neural stem cell proliferation requires CCg, a novel autocrine/paracrine cofactor. J. Biol. Chem. 275, 28369–28375 CrossRef Medline
19. Schüttelkopf, A. W., Hamilton, G., Watts, C., and van Aalten, D. M. (2006) Structural basis of reduction-dependent activation of human cystatin F. J. Biol. Chem. 281, 16570–16575 CrossRef Medline
20. Sotiropoulou, G., Anisowicz, A., and Sager, R. (1997) Identification, cloning, and characterization of cystatin M, a novel cysteine proteinase inhibitor, down-regulated in breast cancer. J. Biol. Chem. 272, 903–910 CrossRef Medline
21. Grubbs, A. O. (2000) Cystatin C: properties and use as diagnostic marker. Adv. Clin. Chem. 35, 63–99 Medline
22. Cheng, T., van Vlijmen-Willems, I. M., Hitomi, K., Pasch, M. C., van Erp, P. E., Schalkwijk, J., and Zeeuwen, P. L. (2009) Colocalization of cystatin M/E and its target proteases suggests a role in terminal differentiation of human hair follicle and nail. J. Invest. Dermatol. 129, 1232–1242 CrossRef Medline
23. Abrahamsson, M., and Grubbs, A. (1994) Increased body temperature accelerates aggregation of the Leu-68→Gln mutant cystatin C, the amyloid-forming protein in hereditary cystatin C amyloid angiopathy. Proc. Natl. Acad. Sci. U.S.A. 91, 1416–1420 CrossRef Medline
24. Ekiel, I., and Abrahamsson, M. (1996) Folding-related dimerization of human cystatin C. J. Biol. Chem. 271, 1314–1321 CrossRef Medline
25. Janowski, R., Abrahamsson, M., Grubbs, A., and Jaskolski, M. (2004) Domain swapping in N-truncated human cystatin C. J. Mol. Biol. 341, 151–160 CrossRef Medline
26. Janowski, R., Kozak, M., Jankowska, E., Grzonka, Z., Grubbs, A., Abrahamsson, M., and Jaskolski, M. (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. Nat. Struct. Biol. 8, 316–320 CrossRef Medline
27. Orlikowska, M., Jankowska, E., Kołodziejczyk, R., Jaskolski, M., and Szymańska, A. (2011) Hinge-loop mutation can be used to control 3D domain swapping and amyloidogenesis of human cystatin C. J. Struct. Biol. 173, 406–413 CrossRef Medline
28. Ghiso, J., Jenson, O., and Frangiote, B. (1986) Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Icelandic type is a variant of γ-trace basic protein (cystatin C). Proc. Natl. Acad. Sci. U.S.A. 83, 2974–2978 CrossRef Medline
29. Yamada, M. (2000) Cerebral amyloid angiopathy: an overview. Neuropathology 20, 8–22 CrossRef Medline
30. Zerovnik, E., Pompe-Novak, M., Skarabot, M., Ravnikar, M., Musevic, I., and Turk, V. (2002) Human legumain B readily forms amyloid fibrils in vitro. Biochim. Biophys. Acta 1594, 1–5 CrossRef Medline
31. Zerovnik, E., Turk, V., and Waltho, J. P. (2002) Amyloid fibril formation by human legumain B: influence of the initial pH-induced intermediate state. Biochem. Soc. Trans. 30, 543–547 CrossRef Medline
32. Skerget, K., Taler-Verbic, A., Badvek, A., Hodnik, V., Ceru, S., Tusek-Znidaric, M., Kumm, T., Pitsi, D., Pompe-Novak, M., Palumaa, P., Soriano, S., Kopitar-Jerala, N., Turk, V., Anderluh, G., and Zerovnik, E. (2010) Interaction between oligomers of legumain B and amyloid-β in vitro and in cells. J. Biol. Chem. 285, 3201–3210 CrossRef Medline
33. Briggs, J. J., Haugen, M. H., Johansen, H. T., Riker, A. I., Abrahamsson, M., Fodstad, Ø., Maaløe, G. M., and Solberg, R. (2010) Cystatin E/M suppresses legumain activity and invasion of human melanoma. BMC Cancer 10, 17 CrossRef Medline
34. Vigneswaran, N., Wu, J., Nagaraj, N., James, R., Zeeuwen, P., and Zacharias, W. (2006) Silencing of cystatin M in metastatic oral cancer cell line MDA-868LN by siRNA increases cysteine proteinases and legumain activities, cell proliferation and in vitro invasion. Life Sci. 78, 898–907 CrossRef Medline
35. Janowski, R., Koza, M., Jankowska, E., Grzonka, Z., Grubbs, A., and Abrahamsson, M. (2004) Do-
Cystatin E transforms to multimers with distinct functions

Amyloid-forming structural superfamily. EMBO J. 20, 4774–4781 CrossRef Medline

Knowles, T. P., Vendruscolo, M., and Dobson, C. M. (2014) The amyloid state and its association with protein misfolding diseases. Nat. Rev. Mol. Cell Biol. 15, 384–396 CrossRef Medline

O’Brien, R. J., and Wong, P. C. (2011) Amyloid precursor protein processing and Alzheimer’s disease. Annu. Rev. Neurosci. 34, 185–204 CrossRef Medline

Mi, W., Pawlik, M., Sastre, M., Jung, S. S., Radinsky, D. S., Klein, A. M., Sommer, J., Schmidt, S. D., Nixon, R. A., Mathews, P. M., and Levy, E. (2007) Cystatin C inhibits amyloid-β deposition in Alzheimer’s disease mouse models. Nat. Genet. 39, 1440–1442 CrossRef Medline

Nguyen, A., and Hulleman, J. D. (2016) Evidence of alternative cystatin C signal sequence cleavage which is influenced by the A25T polymorphism. PLoS One 11, e0147684 CrossRef Medline

Yamane, T., Takeuchi, K., Yamamoto, Y., Li, Y. H., Fujiwara, M., Nishi, K., Takahashi, S., and Ohkubo, I. (2002) Legumain from bovine kidney: its purification, molecular cloning, immunohistochemical localization and degradation of annexin II and vitamin D-binding protein. Biochim. Biophys. Acta 1596, 108–120 CrossRef Medline

Dall, E., and Brandstetter, H. (2012) Activation of legumain involves proteolytic and conformational events, resulting in a context- and substrate-dependent activity profile. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 68, 24–31 CrossRef Medline

Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N., and Nordlund, P. (2006) Thermostabilization of proteins for structural studies. Anal. Biochem. 357, 289–298 CrossRef Medline

Niesen, F. (2010) Excel Script for the Analysis of Protein Unfolding Data Acquired by Differential Scanning Fluorimetry (DSF), version 3.0. Structural Genomics Consortium, Oxford

Williams, J. W., and Morrison, J. F. (1979) The kinetics of reversible tight-binding inhibition. Methods Enzymol. 63, 437–467 CrossRef Medline

Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. D Biol. Crystallogr. 67, 271–281 CrossRef Medline

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242 CrossRef Medline

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 CrossRef Medline

Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 CrossRef Medline

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. 68, 352–367 CrossRef Medline

Delano, W. L. (2002) The PyMOL Molecular Graphics System, version 1.5.0.4, Schroedinger, LLC, New York

Sippl, M. J. (2008) On distance and similarity in fold space. Bioinformatics 24, 872–873 CrossRef Medline

Lo, M. C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., and Elledst, G. (2004) Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. Anal. Biochem. 332, 153–159 CrossRef Medline