Cathepsin D Propeptide: Mechanism and Regulation of Its Interaction with the Catalytic Core†

Martin Máša, Lucie Marešová, Jiří Vondrášek, Martin Horn, Jan Ježek, and Michael Mareš*

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, 16610 Prague, Czech Republic

Received July 25, 2006; Revised Manuscript Received October 26, 2006

ABSTRACT: Propeptide blocks the active site in the inactive zymogen of cathepsin D and is cleaved off during zymogen activation. We have designed a set of peptidic fragments derived from the propeptide structure and evaluated their inhibitory potency against mature cathepsin D using a kinetic assay. Our mapping of the cathepsin D propeptide indicated two domains in the propeptide involved in the inhibitory interaction with the enzyme core: the active site "anchor" domain and the N-terminus of the propeptide. The latter plays a dominant role in propeptide inhibition (nanomolar K_i), and its high-affinity binding was corroborated by fluorescence polarization measurements. In addition to the inhibitory domains of propeptide, a fragment derived from the N-terminus of mature cathepsin D displayed inhibition. This finding supports its proposed regulatory function. The interaction mechanisms of the identified inhibitory domains were characterized by determining their modes of inhibition as well as by spatial modeling of the propeptide in the zymogen molecule. The inhibitory interaction of the N-terminal propeptide domain was abolished in the presence of sulfated polysaccharides, which interact with basic propeptide residues. The inhibitory potency of the active site anchor domain was affected by the Ala38pVal substitution, a propeptide polymorphism reported to be associated with the pathology of Alzheimer's disease. We infer that propeptide is a sensitive tethered ligand that allows for complex modulation of cathepsin D zymogen activation.

Pepsin-like aspartic proteases are synthesized at neutral pH as inactive zymogens that are inhibited by the binding of their N-terminal propeptides in the active site. The zymogens can be autoactivated to their mature forms upon acidification, which leads to a concerted series of conformational changes concomitant with both intra- and intermolecular processing events (for a review, see ref *1*). The proteolytic removal of the propeptide (so-called activation peptide) proceeds in one or more steps, and the variability of this mechanism of proteolysis appears to be attributable to the structure of the prosegment as well as to environmental factors. In 1978, Dunn et al. found that propeptide fragments released during activation of pepsinogen can function as inhibitors of mature pepsin (*2*). Subsequently, this phenomenon was demonstrated for fragments derived from the propeptide of renin (*3*, *4*). We have previously reported that the full-length propeptides of cathepsin D and pepsin are able to interact with nonparental aspartic proteases, which reflects the sequence similarity of propeptides and suggests common putative interaction motifs (*5*, *6*). Recognition of the propeptide by the mature enzyme was applied for isolation of cathepsin D by affinity chromatography (*7*).

Cathepsin D, a lysosomal aspartic protease, is involved in the turnover of cellular proteins as well as in the selective processing of MHC II antigens, hormones, and growth factors. Gene knockout studies have demonstrated that cathepsin D-deficient mice die prematurely from massive physiological abnormalities that indicate misregulated cell growth and tissue homeostasis (*8*). Recently, a cathepsin D deficiency caused by gene mutations has been shown to cause congenital neuronal ceroid-lipofuscinosis (*9*, *10*). A polymorphism in the cathepsin D gene, which results in the Ala38pVal substitution in the propeptide, was reported to be associated with the pathogenesis of Alzheimer's disease (*11*, *12*). Cathepsin D is an independent marker of poor prognosis in human breast cancer (*13*). Overexpression and secretion of cathepsin D zymogen have been implicated in progressive breast and prostate cancer: the zymogen functions as a mitogen that interacts, possibly through its propeptide, with an unknown receptor on cancer cells (*14*, *15*). Cathepsin D is also a key mediator of induced apoptosis in the tumor cell cytosol (*16*).

The X-ray structure of mature cathepsin D obtained from crystals grown at acidic pH shows a protein fold common to aspartic proteases (*17*, *18*). However, mature cathepsin D crystallized at pH 7.5 was found to be structurally rearranged in a catalytically inactive form, in which the N-terminal strand relocates and inserts into the active site cleft (*19*). This conformation is believed to correspond to a hypothetical transient "intermediate 3" formed during the activation cascade of pepsin-like protease zymogens. The activation "intermediate 2", which contains the noncovalently bound portion of the cleaved propeptide, has been crystallographically characterized for the aspartic proteases gastricsin and cathepsin E (*20*, *21*). The intact cathepsin D zymogen has not yet been crystallized, but zymogen structures of its closest

† Supported by MSMTCR Grant 1P04OCB20.002 and COST B20 Action, GAAVCR Grant IAA4055303, and Research Project Z40550506.

* To whom correspondence should be addressed. Phone: (420) 220183358. Fax: (420) 220183578. E-mail: mares@uochb.cas.cz.

10.1021/bi0614986 CCC: $33.50
Published on Web 12/08/2006

mammalian homologues pepsinogen and progastricsin are available (*22*, *23*).

In this work, we present the first detailed, systematic mapping of cathepsin D propeptide to localize the structural determinants involved in propeptide−enzyme recognition. We have identified three autonomous inhibitory domains in the propeptide and adjacent mature N-terminus responsible for inactivation of the enzyme core. The regulatory function of these domains is interpreted with the help of the homology model of the cathepsin D zymogen. The uncovered inhibitory motifs represent potential leads for designing a new class of inhibitors of cathepsin D and other aspartic proteases. Finally, we propose novel physiologically relevant mechanisms of modulation of the activation process of the cathepsin D zymogen through propeptide mutation and interaction of the propeptide with glycosaminoglycans.

EXPERIMENTAL PROCEDURES

Materials. Mature cathepsin D was purified to homogeneity from human placenta according to Fusek et al. (*24*). Heparin, dextran sulfate, and 5-(iodoacetamido)fluorescein were from Sigma (St. Louis, MO), and pepstatin A was from Fluka (Buchs, Switzerland). The fluorogenic substrate Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu (Abz,[1] aminobenzoic acid; Nph, 4-nitrophenylalanine) was designed on the basis of a chromogenic cathepsin D substrate and equipped with an Abz group for FRET (fluorescence resonance energy transfer) application (*5*). Amino acid precursors were obtained from NovaBiochem (Darmstadt, Germany); resins and other substances for peptide synthesis were from Bachem (Bubendorf, Switzerland).

Synthesis of Peptides. The peptides were synthesized by Fmoc solid phase chemistry on an ABI 433A peptide synthesizer (Applied Biosystems, Foster City, CA). The peptides composed of L-amino acids were prepared as peptidyl amides with free N-termini. The fluorescein-labeled derivative of peptide [1p−14p] was synthesized with an additional C-terminal cysteine residue that was postsynthetically modified with 5-(iodoacetamido)fluorescein according to the manufacturer's protocol (Sigma). All peptides were purified by reverse phase HPLC on a Vydac C_{18} column (218TP510, Vydac, Hesperia, CA) equilibrated in 0.1% (v/v) TFA and eluted with a 1%/min gradient of a 90% (v/v) acetonitrile solution in 0.1% (v/v) TFA. The purified peptides were characterized by ESI mass spectrometry on an LCQ Classic Finnigan Mat device (Thermo Finnigan, Bremen, Germany).

Inhibition Assay. Inhibition of cathepsin D activity by synthetic peptides was assayed with the fluorogenic substrate Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu. The reaction was performed in microplate format in a total assay volume of 100 μL. The preincubation mixture (50 μL) contained 0−300 μM peptide and 1.2 nM cathepsin D (final concentrations in the assay) in 25 mM Tris-HCl buffer (pH 7.5). The mixture was incubated at 37 °C for 10 min, followed by the addition of 50 μL of substrate (35 μM in the assay) in 0.3 M sodium acetate buffer (pH 5.1) (or pH 3.5−5.5 for analysis of the pH dependence) to set the final pH. Where indicated, the sulfated polysaccharides were added to the preincubation mixture to give a final concentration of 0−5 μg/mL in the assay. The kinetics of product release were continuously monitored on a GENios microplate reader (TECAN, Salzburg, Austria) at excitation and emission wavelengths of 330 and 410 nm, respectively. Each measurement was performed in triplicate. The inhibition constants (K_i) were determined from the residual velocities using the dose−response plot (v_i/v_0 vs [I]) to obtain IC_{50} values that were converted to K_i values by Cheng and Prusoff relationships (*25*). The mechanism of inhibition was determined using an analogous activity assay but performed at a constant pH of 5.1 under steady-state conditions; the initial velocities of product release were analyzed with a Lineweaver−Burk plot. The stock concentration of cathepsin D was determined by titration with pepstatin A; the peptide solutions were quantified by amino acid analysis.

Stability of Peptides. A reaction mixture, composed of 60 μM peptide and 10 nM cathepsin D, was incubated in 100 mM sodium acetate buffer (pH 5.1) at 37 °C for 16 h. The reaction was stopped by adding 15% TFA, and the mixture was separated using RP-HPLC. The chromatography was performed on a C_4 Vydac column (214TP54, Vydac) equilibrated in 0.1% (v/v) TFA and eluted with a 1%/min gradient of a 99% (v/v) acetonitrile solution in 0.1% (v/v) TFA. The collected peak fractions were analyzed by ESI mass spectrometry (LCQ Classic Finnigan Mat).

Secondary Structure Analysis. The circular dichroism spectra were measured with 0.5 mM peptide in 10 mM sodium phosphate (pH 7) containing 50% (v/v) trifluoroethanol at 26 °C using a CD6 Dichrograph (Jobin Yvon, Longjumea, France). Prediction of secondary structure for the cathepsin D propeptide was performed by the neural networks method run by the PHD/PROF program (http://www.predictprotein.org).

Molecular Modeling. The spatial model of the cathepsin D zymogen was built using the X-ray structure of mature cathepsin D (PDB entry 1LYA) as a template of the enzyme core. The cathepsin D propeptide (GenBank accession number P07339) structure was modeled using the X-ray structure of pepsinogen propeptide (PDB entry 2PSG) as a template according to the structure-based alignment in Figure 1. Molecular replacement of the structurally conserved regions and rebuilding of the variable regions were conducted with the homology module of MOE (Chemical Computing Group Inc., Montreal, PQ). The structure was subjected to minimization (convergence criterion of 0.001 kcal mol^{-1} $Å^{-1}$) using the Cornell force field (*26*). The root-mean-square deviation of the framework (C_α) was ∼2.4 Å in the energy-optimized model of the cathepsin D zymogen as compared to pepsinogen.

Fluorescence Measurements. The fluorescence polarization measurements were performed with an Aminco-Bowman Series 2 luminescence spectrometer (Thermo Electron, Cambridge, U.K.) at 25 °C using excitation and emission wavelengths of 485 and 530 nm, respectively. Anisotropy titration was carried out in competition experiments with 0−6 μM unlabeled peptide [1p−14p], 20 nM fluorescein-labeled peptide [1p−14p] as a "reporter ligand", and 5 nM cathepsin D in 100 mM Tris-HCl buffer (pH 7.5). The binding mixture was incubated for 1 h at 35 °C to reach equilibrium. The fluorescence of the buffer alone was subtracted from each

[1] Abbreviations: Abz, aminobenzoic acid; Nph, 4-nitrophenylalanine; TFA, trifluoroacetic acid.

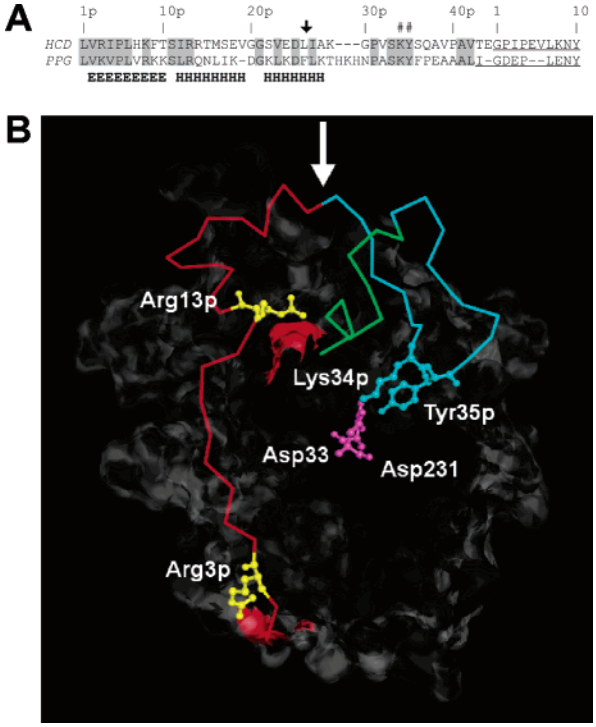

FIGURE 1: (A) Structure-based alignment of propeptides of human cathepsin D (HCD) and pepsinogen (PPG) (GenBank accession numbers P07339 and P00791, respectively). The consensus secondary structure is based on X-ray data of PPG (PDB entry 3PSG) and prediction analysis of HCD: H, α-helix; E, *β*-strand. Anchor residues Lys34p and Tyr35p interact with the catalytic residues of the enzyme core (#); the position of the autoprocessing site in the cathepsin D propeptide (arrow) and mature N-termini (underlined) are indicated. Homologous residues are highlighted in gray. The numbering is according to cathepsin D; the suffix p indicates propeptide numbering. (B) The tertiary structure of procathepsin D was modeled using the X-ray structures of mature cathepsin D (PDB entry 1LYA) and pepsinogen (PDB entry 3PSG). The enzyme core is shown as a white surface (transparent at the front). The propeptide (backbone) contains two major regions used for the design of the propeptide fragments (Figure 2): the N-terminal portion (residues 1p−26p, red) and the C-terminal portion (residues 27p−44p, cyan) separated by the autoprocessing site (arrow). Ten green-colored residues adjacent to the propeptide correspond to the mature N-terminus. The following interactions are indicated: propeptide residues Arg3p and Arg13p (yellow sticks) with core residues Asp181 and Asp12 (red surface), respectively, and propeptide anchor residues Lys34p and Tyr35p (blue sticks) with catalytic Asp33 and Asp231 residues (magenta sticks). This figure was prepared with UCSF Chimera (*28*).

measurement, and the anisotropy was calculated according to ref *27*. *G* factor correction was performed using fluorescein as a standard. The anisotropy data were converted to the fraction of the reporter ligand bound using the following equation:

$$f_b = \frac{r_{obs} - r_0}{(r_{obs} - r_0) + R(r_{max} - r_{obs})}$$

where f_b is the fraction of the reporter ligand bound and r_{obs}, r_0, and r_{max} are the observed anisotropy, anisotropy of the free reporter ligand, and anisotropy of the reporter ligand with cathepsin D, respectively. *R* is the fluorescence intensity of the reporter ligand with cathepsin D divided by the intensity of the free reporter ligand. The K_d of the binding of unlabeled ligand to cathepsin D was determined by a nonlinear least-squares fit to a single-site binding model using Grafit (Erithacus software, Surrey, U.K.). The stock concentration of cathepsin D was determined by titration with pepstatin A; the peptide solutions were quantified by amino acid analysis.

RESULTS

Design of Propeptide Fragments. We constructed a spatial model of human cathepsin D zymogen (procathepsin D) as a tool for structure-based design of peptidic fragments derived from its propeptide. The amino acid sequence of the propeptide of procathepsin D is substantially homologous with those of the mammalian members of the pepsin family, especially gastric aspartic proteases such as pepsinogen (*6*) (Figure 1A). The prediction of secondary structure elements of the procathepsin D propeptide revealed that the general pattern, identified experimentally by X-ray structural analysis in pepsinogen, is preserved (*22*) (Figure 1A). This similarity implies an analogous fold in the propeptide of procathepsin D. The X-ray structure of the pepsinogen propeptide was used as a template for homology modeling of procathepsin D (Figure 1B) (see Experimental Procedures). The model of procathepsin D propeptide consists of an N-terminal *β*-strand followed by a centrally positioned α-helix. An autoprocessing site is located in the center of the α-helix, and its cleavage results in the formation of an active semiprocessed form with a shortened propeptide called pseudocathepsin D (*29*, *30*). Downstream of the surface-exposed α-helix, the propeptide turns to the bottom of the active site. The conserved residues Lys34p and Tyr35p in propeptide and Tyr8 in the mature N-terminal region interact directly with the catalytic aspartates and thus stabilize the inactive zymogen structure by an internal anchor. We suppose that the remaining C-terminal part of the propeptide is exposed, has little secondary structure, and may exhibit conformational flexibility.

The list of peptide fragments derived from the sequence of propeptide as well as from the N-terminal sequence of mature human cathepsin D is shown in Figure 2. Two large fragments, [1p−26p] and [27p−44p], span the entire 44-residue sequence of propeptide; the Leu26p−Ile27p bond located at the boundary forms the autoprocessing site. The overlapping fragment, [15p−36p], corresponds to the central portion of propeptide that adopts an α-helical conformation in the procathepsin D model. Fragment [1p−26p] was gradually truncated at its C-terminus to produce shorter fragments of 14 and 10 residues containing the N-terminal *β*-strand. The C-terminal stretch of propeptide [27p−44p] and the following 10 residues of mature N-terminus were scanned by a set of overlapping fragments. The modeled propeptide segment around anchor residues Lys34p and Tyr35p forms a compact structure embedded in the catalytic site; this fragment was synthesized in linear conformation as well as constrained conformation using disulfide bridging.

Mapping of Inhibitory Domains. The synthetic propeptide fragments were assayed for their ability to inhibit the enzymatic activity of mature human cathepsin D using a kinetic assay with a fluorogenic substrate. The peptides were preincubated with enzyme at pH 7.5 to allow formation of an inhibitory complex with a conformation similar to that in procathepsin D, which is stabilized at neutral pH (*5*). The

FIGURE 2: Design and inhibitory activity of fragments derived from the propeptide of human cathepsin D. The sequence of the propeptide (residues 1p−44p) is shown together with the N-terminal sequence of mature cathepsin D (highlighted in gray). Anchor residues Lys34p and Tyr35p predicted to interact with the catalytic residues of the enzyme core are marked (#); the arrow shows the position of the autoprocessing site. Synthetic peptide fragments are coded according to their location in the sequence. The substituted residues are in bold and underlined. B denotes citrulline; the Ala38pVal substitution corresponds to cathepsin D gene polymorphism. Disulfide bridging in peptide [31p−38p]Cys-Cys is indicated. The peptides were screened for their potential to inhibit the activity of mature cathepsin D, and their inhibition constant (K_i) is given. Mean values ± the standard error are given for triplicate measurements. One asterisk means the value is reported in ref *5*; two asterisks mean the measurement of inhibition is influenced by degradation of the peptide inhibitor. NI indicates no significant inhibition with 300 μM peptide.

substrate reaction was initiated by lowering the pH to 5.1, at which procathepsin D is activated and the mature cathepsin D is enzymatically active (*29*−*31*). The inhibition constants (K_i) of the fragments delineate individual structural domains possessing an intrinsic inhibitory potency (Figure 2). The mapping showed that the most inhibitory domain is located at the N-terminus of the propeptide. All fragments derived from this domain exhibited submicromolar K_i values. The lowest K_i (∼10 nM) was found for 14-residue fragment [1p−14p] and was significantly increased by shortening the peptide length to 10 residues (fragment [1p−10p]). This suggests that Arg13p, which forms an electrostatic interaction with Asp12 of the enzyme core (Figure 1B), is important for efficient inhibition. The 12 subsequent residues do not improve inhibition, as demonstrated by the longer fragment, [1p−26p].

The fragments derived from the C-terminal stretch of propeptide (downstream of the Leu26p−Ile27p autoprocessing site) displayed K_i values in the two-digit micromolar range and above. The data with the shortened and substituted peptides demonstrate that a sufficient fragment length is important for the inhibition and that anchor residues Lys34p and Tyr35p are among the required structural determinants (Figure 2). This points to the fact that conformational support of the propeptide scaffold is necessary for the establishment of an effective inhibitory interaction at the catalytic site. The inhibitory potency of C-terminal fragments can be significantly improved if they are extended to include the sequence of the mature N-terminus. We found that such potentiation of inhibition is associated with residues 5−10 (Figure 2). Moreover, the mature N-terminal fragment [1−10] itself functioned as an autonomous inhibitory domain with a K_i of ∼28 μM. Its inhibitory potency was abolished when Lys8 and Tyr10 were mutated. These residues represent important stabilization residues in mature cathepsin D for the conformations at acidic as well as neutral pH (*17*−*19*).

Proteolysis of Propeptide Fragments. The resistance of propeptide-derived fragments to proteolysis was investigated under conditions similar to those used in the inhibition assay. The incubation mixture of the peptides with mature cathepsin D was analyzed by analytical RP-HPLC, and the resulting profile was compared to that of the nondigested control. The peptides exhibited stability against proteolysis at pH 5.1 with the exception of fragment [15p−36p], which spans the central part of the propeptide. Fragment [15p−36p] was specifically cleaved at the Leu26p−Ile27p bond (Figure 3A). This position is identical to the autoprocessing site that is split during the acidic autoactivation of procathepsin D (*29*−*31*).

In the procathepsin D model, fragment [15p−36p] covers the helical region of the propeptide (Figure 1B). We therefore directly investigated the secondary structure of this synthetic fragment by circular dichroism spectroscopy (Figure 3B).

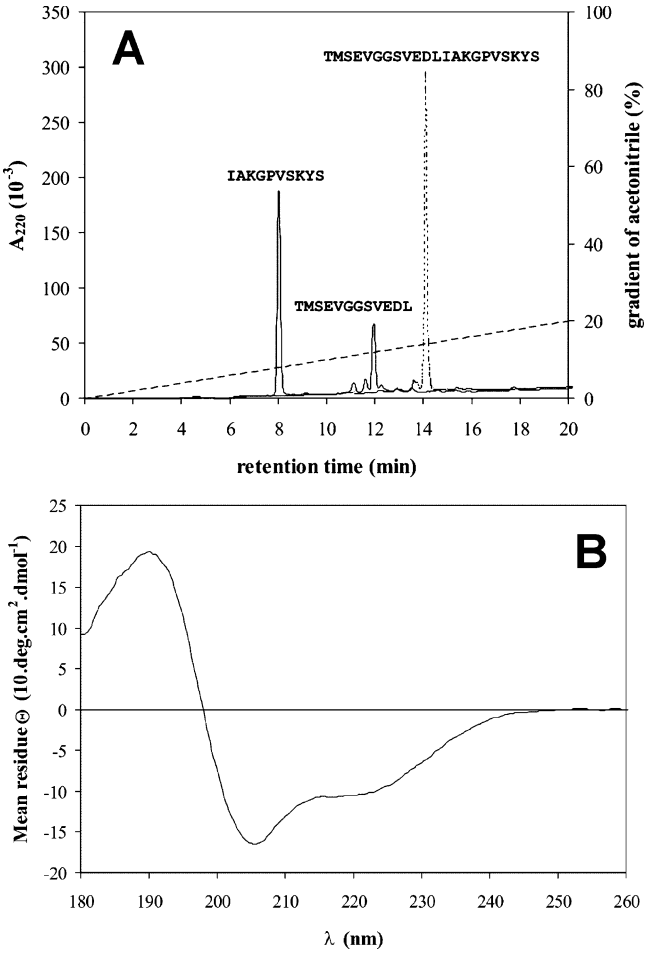

FIGURE 3: Analysis of central fragment [15p−36p] derived from the cathepsin D propeptide. (A) RP-HPLC profile of synthetic peptide [15p−36p] (···) and its digest by mature cathepsin D performed at pH 5.1 (—). The digestion products resulting from cleavage of the Leu26p−Ile27p bond were identified by ESI mass spectrometry; the peptide sequences are indicated. Chromatography was performed on a C_4 Vydac column equilibrated in 0.1% (v/v) TFA and eluted with a 1%/min gradient of 99% (v/v) acetonitrile (- - -). (B) The circular dichroism spectrum of peptide [15p−36p] indicates an α-helical conformation. The measurement was performed in the presence of 50% trifluoroethanol at pH 7.5.

The spectrum implied a partial α-helical conformation amplified in the presence of trifluoroethanol (*32*). This finding suggests that the free propeptide fragment tends to adopt the same secondary structure as the corresponding region in the full-length propeptide bound in the zymogen molecule. This conformational similarity likely contributes to the identity of the cleavage site in both the fragment and the whole zymogen.

Type of Inhibition. A detailed kinetic analysis was performed to determine the mode of inhibition of mature cathepsin D by the discovered inhibitory fragments. We selected fragments [1p−14p], [27p−44p], and [1−10] as representatives of the three localized inhibitory domains, namely, the N-terminus of the propeptide, the C-terminal anchor of the propeptide, and the mature N-terminus, respectively (Figure 2). The hydrolysis of the fluorogenic peptidic substrate at pH 5.1 follows standard Michaelis−Menten kinetics with a K_m of 33 μM. The initial rates of product formation in the presence of inhibitory fragments are shown in a Lineweaver−Burk plot (Figure 4). For fragment [27p−44p], the family of straight lines intersects on the $1/v$ axis, indicating competitive inhibition ($K_i \sim 9$ μM). For fragments [1p−14p] ($K_i \sim 11$ nM) and [1−10] ($K_i \sim 198$ μM), the kinetic analysis is consistent with a noncompetitive mode of inhibition as indicated by the position of the intercept point on the 1/[S] axis. The inhibitory potencies determined here for the propeptide-derived domains ([1p−14p] and [27p−44p]) are on the same order of magnitude as those measured after preincubation with the enzyme at pH 7.5 prior to the assay at pH 5.1 (Figure 2). This is in contrast with the full-length propeptide which displayed inhibition 3 orders of magnitude stronger when assayed after preincubation at neutral pH (*5*). The fact that the inhibition of the full-length propeptide is more sensitive to the pH environment than the individual inhibitory domains may reflect a cooperative conformational behavior of the full-length ligand during its interaction.

Fluorescence Binding Study. Fluorescence polarization experiments were used to determine the equilibrium dissociation constant for the complex of mature cathepsin D with fragment [1p−14p], which spans the most inhibitory domain (Figure 2). This approach allows for direct measurement of its affinity at neutral pH, at which cathepsin D is proteolytically inactive and the zymogen-like conformation is stabilized. For this purpose, we synthesized a fluorescein-labeled derivative of [1p−14p] with a C-terminally attached fluorescein moiety to preserve the native N-terminal amino group of the fragment (see Experimental Procedures). This derivative has inhibitory potency ($K_i \sim 16$ nM at pH 5.1) within the same order of magnitude as the original fragment, [1p−14p] [$K_i \sim 10$ nM (Figure 2)]. The labeled derivative was used as a ligand in a saturation binding experiment monitored by fluorescence anisotropy that showed a typical titration curve of the labeled [1p−14p] peptide against the constant concentration of the enzyme (data not shown). The fluorescence anisotropy measurement was then applied in a competition assay between labeled and unlabeled [1p−14p] peptide to yield a K_d of ~89 nM for the original fragment, [1p−14p] (Figure 5).

Arg Mutations in Propeptide. According to the procathepsin D model, the N-terminal segment of propeptide forms electrostatic interactions with the enzyme core (Figure 1B). In particular, Arg3p and Arg13p, basic residues also conserved in the pepsinogen propeptide, interact with core residues Asp181 and Asp12, respectively. We investigated how substitution of the arginine residues influences the inhibitory potential of the most inhibitory fragment, [1p−14p]. Derivatives of [1p−14p] in which Arg3p and Arg13p were replaced with citrulline, glutamic acid, or leucine were synthesized and tested in the activity assay with mature cathepsin D (Figure 2). The substitution of arginines with isosteric citrulline with an uncharged side chain ([1p−14p]-Cit_2) led to a significant decrease in the inhibitory potency; substitution with the oppositely charged residues of glutamic acid ([1p−14p]Glu_2) resulted in a profound, 2-order-of-magnitude increase in K_i. These results clearly point to the stabilizing role of ionic interactions. On the other hand, "hydrophobized" derivative [1p−14p]Leu_2 displays a K_i lower than that of [1p−14p]Cit_2. This suggests that hydrophobic interactions, which are plentiful in the N-terminal β-strand of the procathepsin D model, are more important than ionic interactions. Furthermore, we investigated the pH dependence of the inhibitory interaction of [1p−14p] and

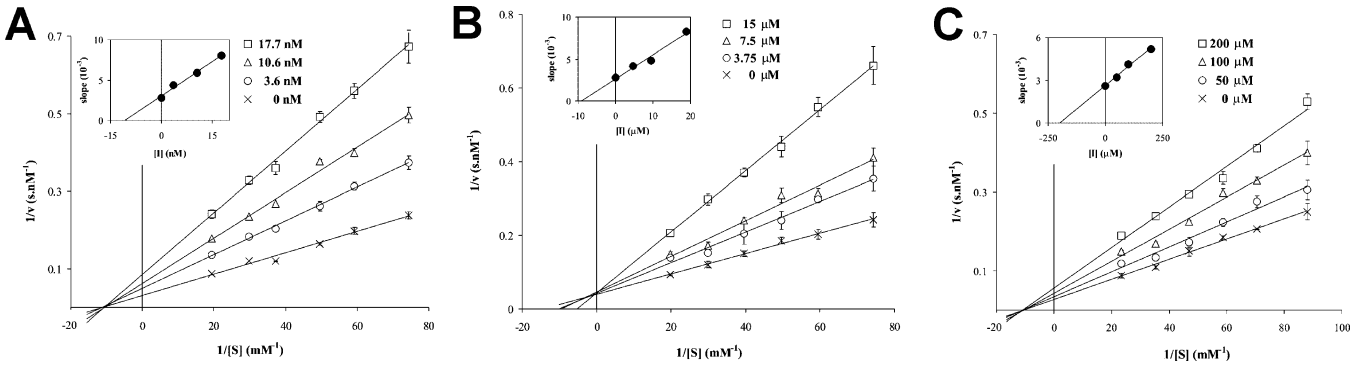

FIGURE 4: Mode of inhibition of mature cathepsin D by fragments derived from the inhibitory domains. Lineweaver−Burk plots are presented together with secondary plots of the same data (inset). Fragments [1p−14p] from propeptide (A) and [1−10] from the mature N-terminus (C) display noncompetitive inhibition, and fragment [27p−44p] from propeptide (B) displays competitive inhibition. The assay was performed at pH 5.1. Error bars depict the standard deviation of the mean of triplicate measurements.

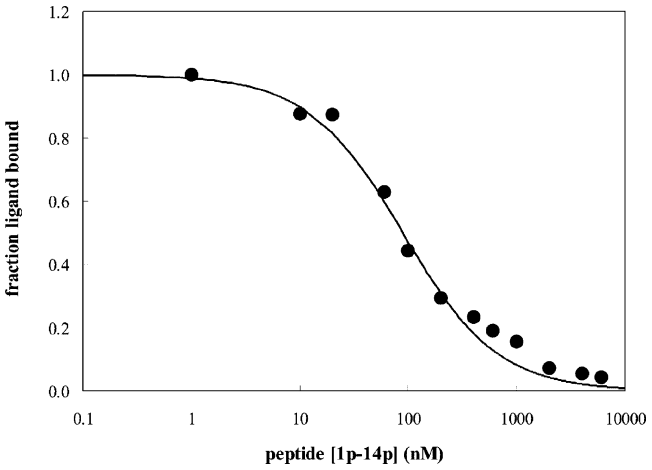

FIGURE 5: Fluorescence polarization assay of the binding of fragment [1p−14p] to mature cathepsin D at pH 7.5 plotted as a fraction of the ligand bound. Competitive displacement of the fluorescein-labeled [1p−14p] fragment (20 nM) complexed with cathepsin D (5 nM) by the unlabeled [1p−14p] fragment (concentration indicated) revealed the optimized parameter ($K_d \sim 89$ nM) of the unlabeled fragment.

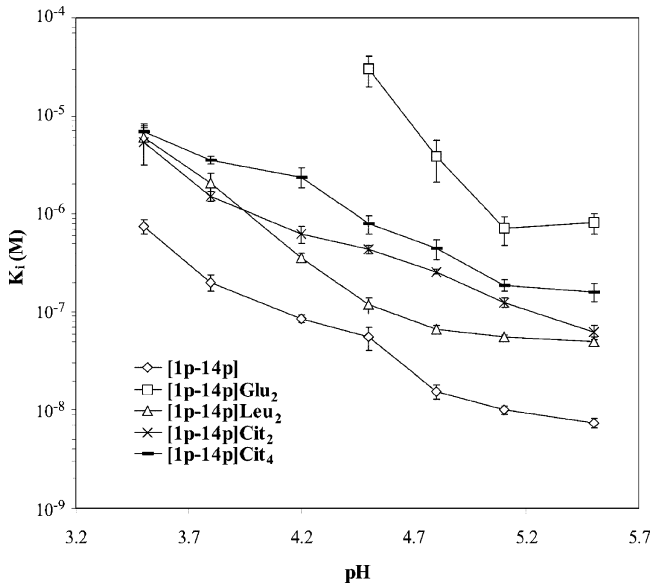

FIGURE 6: pH-dependent inhibitory potency of fragment [1p−14p] and its derivatives. The Arg3p and Arg13p residues in wild-type fragment [1p−14p] were substituted with citrulline ([1p−14p]Cit_2), leucine ([1p−14p]Leu_2), or glutamic acid ([1p−14p]Glu_2). [1p−14p]Cit_4 is substituted in all Arg/Lys positions with citrulline (Figure 2). The peptides were screened for their potential to inhibit the activity of mature cathepsin D. After preincubation of the reaction mixture at pH 7.5, the activity was measured at various pHs to determine the inhibition constant (K_i).

its derivatives. The analysis showed a gradual increase in K_i with a decrease in pH, which is in accord with the activation scheme of pepsin-like protease zymogens based on destabilization of the propeptide under acidic conditions (Figure 6). This general trend is preserved for all derivatives of [1p−14p], including those with substituted Arg residues, which suggests that the effect is caused by pH-dependent changes in the enzyme core, most likely by relocation of the mature N-terminus into the position it occupies at low pH (*33*).

Ala38pVal Mutation. We investigated the physiologically relevant Ala38pVal substitution, which has been reported to be associated with Alzheimer's disease (*11*, *12*). Ala38pVal procathepsin D is a minor genetic variant of the procathepsin D polymorphism (*34*). Residue 38p is localized in the C-terminal inhibitory domain of the propeptide. We synthesized two propeptide fragments spanning residues 27p−44p with an Ala or Val residue at position 38p and tested their inhibitory activity against mature cathepsin D (Figure 2). The peptide corresponding to the major variant Ala38p ([27p−44p]) inhibited with a K_i of ~12.4 μM; however, the Val38p variant ([27p−44p]Val) exhibited no significant inhibition at concentrations up to 300 μM. This difference in inhibitory potential is not due to proteolysis in the assay since neither peptide is significantly degraded by cathepsin D (see Proteolysis of Propeptide Fragments). The impact of the Ala38pVal substitution on the inhibitory interaction may be associated with its location in the spatial vicinity of catalytic site anchor residues Lys34p and Tyr35p, which are critical for inhibition by the C-terminal inhibitory domain of the propeptide (Figure 2).

Interaction with Sulfated Polysaccharides. We analyzed the effect of sulfated polysaccharides (SPs), namely, dextran sulfate and glycosaminoglycan heparin, on the interaction of propeptide with the enzyme core. Fragments [1p−14p] and [27p−44p], covering two inhibitory domains in the propeptide, were tested for inhibition of mature cathepsin D in the presence of SPs. The inhibitory activity of fragment [1p−14p] decreased in the presence of SPs, whereas the inhibitory properties of [27p−44p] remained unchanged. In Figure 7A, we show that modulation by SPs is concentration-dependent with the maximum effect reached at 1 μg/mL SPs,

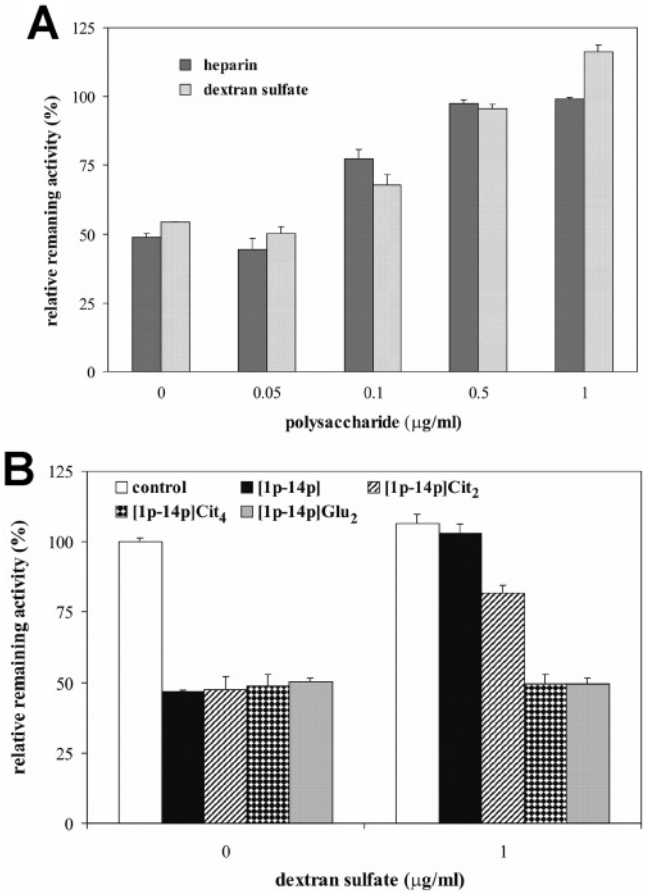

FIGURE 7: Modulatory effect of sulfated polysaccharides on inhibition of mature cathepsin D by fragment [1p−14p]. (A) The inhibitory potency of fragment [1p−14p] is reduced by dextran sulfate and heparin in a dose-dependent manner. (B) The modulatory effect of dextran sulfate is dependent on basic residues in fragment [1p−14p]. The Arg3p and Arg13p residues in wild-type fragment [1p−14p] were substituted with citrulline ([1p−14p]Cit_2) or glutamic acid ([1p−14p]Glu_2). [1p−14p]Cit_4 is substituted in all Arg/Lys positions with citrulline (Figure 2). The reaction mixture of peptide (applied at the K_i concentration), sulfated polysaccharide, and mature cathepsin D was preincubated at pH 7.5; the residual proteolytic activity was measured at pH 5.1 and compared with that of the uninhibited control.

which almost completely abolished the inhibition by [1p−14p]. Furthermore, we investigated the structural basis of the modulatory effect of SPs using the substituted derivatives of [1p−14p] in which lysine/arginine residues were replaced with citrulline or glutamic acid (Figure 7B). The modulatory effect of SPs was weakened by the substitution of two arginine residues with uncharged citrulline ([1p−14p]Cit_2) and was completely eliminated by substitution with oppositely charged glutamic acid ([1p−14p]Glu_2). Also, the substitution of all four basic residues with citrulline made the inhibitory derivative [1p−14p]Cit_4 insensitive to the "anti-inhibitory" effect of SPs. We conclude that the negatively charged SPs interact with the basic residues of propeptide fragment [1p−14p] to obstruct its inhibitory interaction with the enzyme core of cathepsin D.

DISCUSSION

We report the functional mapping of the N-terminal sequence of procathepsin D comprised of 44 propeptide residues and the subsequent 10 residues of the N-terminus of mature cathepsin D. The synthetic peptides derived from this sequence were used as a tool to localize and characterize the structural determinants involved in activity regulation of the enzyme core. The analysis revealed three structural domains capable of establishing the inhibitory interaction: (i) N-terminus of the propeptide, (ii) active site anchor domain of the propeptide, and (iii) N-terminus of the mature sequence.

The N-terminus of the propeptide, which spans residues 1p−14p, displayed the highest inhibitory activity with a K_i of ~10 nM among all fragments tested. This value is fairly comparable to the inhibition strength of the full-length propeptide (Figure 2). As the second inhibitory domain identified in the propeptide (the active site anchor) exhibits only micromolar inhibition, we conclude that the N-terminal propeptide domain controls the inhibitory strength of the entire propeptide. The inhibitory interaction of fragment [1p−14p] with cathepsin D monitored at acidic pH (using the activity assay) is preserved at neutral pH (when the zymogen is inactive) as confirmed by the fluorescence anisotropy assay.

Fragment [1p−14p] obeys noncompetitive inhibition kinetics, which can be interpreted with the help of the putative three-dimensional model. The procathepsin D model suggests that the N-terminal domain of the propeptide adopts a conformation analogous to that observed in activation intermediate 2 of progastricsin (hGSi) (*20*). In the structure of hGSi, propeptide segment 1p−26p is noncovalently bound to the enzyme core in a zymogen-like orientation. It indirectly obstructs the substrate binding cleft by preventing the conformational fine-tuning of the active site to become catalytically efficient. Arg14p in hGSi (corresponding to Arg13p in procathepsin D) has no steric clash with the substrate analogue pepstatin superimposed in the active site (*20*). Such interaction would comply with the noncompetitive inhibition observed for fragment [1p−14p]. In conclusion, we predict that the inhibitory complex of [1p−14p] with cathepsin D mimics intermediate 2. It is stabilized by Arg13p_Asp12 and Arg3p_Asp181 salt bridges [as evidenced by the substituted derivatives of [1p−14p] (Figure 2)] as well as by extensive hydrophobic interactions [through the N-terminal β-strand (Figure 1)]. Furthermore, we investigated whether the interaction of [1p−14p] can be influenced by sulfated polysaccharides (SPs), including glycosaminoglycans. These macromolecules are known to interact with cysteine cathepsins to modulate their proteolytic activity, pH stability, or activation processing (*35*−*38*). The effect of SPs on aspartic cathepsins, however, has not been published to date. We discovered that SPs efficiently abolish the inhibition of cathepsin D by [1p−14p]. The mechanism of this anti-inhibitory effect is based on electrostatic interactions between SPs and basically charged residues of [1p−14p]. We suppose that SPs represent physiologically relevant compounds that may weaken the affinity of the propeptide for zymogen and hence accelerate in vivo activation of procathepsin D. Experimental evaluation of this hypothesis is currently underway.

The second inhibitory domain we identified in the propeptide surrounds Lys34p and Tyr35, and these two residues are essential for binding of the domain. The domain topology in the procathepsin D model explains its competitive inhibition. The Lys34p-Tyr35p pair functions as an anchor that

directly interacts with the catalytic aspartates, analogous to the interaction in pepsinogen or progastricsin (*22*, *23*). The adjacent residues flanking the anchor generally have a smaller number of interactions with the core. Accordingly, we found that only larger peptides containing the Lys34p-Tyr35p pair provide an effective scaffold that is able to establish the interaction (Figure 2). The inhibition is nevertheless relatively weak (K_i in micromolar range), making this domain sensitive to environment and potential regulation. Thus, the disruption of the salt bridges of the Lys34p-Tyr35p pair upon acidification will destabilize the entire region and initiate the activation process common for pepsinogen-like zymogens (*20*). Finally, we analyzed whether the Ala38pVal substitution located in this domain has any functional consequence. This substitution was originally identified as a missense polymorphism in the procathepsin D gene (*34*). The Ala derivative (major variant) of the synthetic domain exhibited a K_i of ~12 μM, in clear contrast to the Val derivative (minor variant), which did not exhibit any inhibition (Figure 2). This indicates that local structural changes in the vicinity of the Lys34p-Tyr35p pair can critically modulate properties of this propeptide domain. The finding is relevant from a genetic epidemiology standpoint as several studies reported that this polymorphism is associated, through an unknown mechanism, with the pathogenesis of Alzheimer's disease (*11*, *12*). Here, we provide for the first time experimental data which show that Ala38pVal polymorphism may modulate the propeptide interaction and hence the procathepsin D activation.

The inhibitory domain we localized in the N-terminus of mature cathepsin D proves that there is a tethered structure outside of the propeptide functioning in autoregulation of the enzyme core. At low pH, the active site of cathepsin D is freely accessible, and the mature N-terminus takes the position originally occupied by the N-terminus of the propeptide in the zymogen molecule (*18*). At pH 7.5, the mature N-terminus is repositioned into the active site cleft and interacts with the catalytic residues, the mechanism of which was proposed to function as a pH-dependent activity regulation (*19*). It has been demonstrated recently that the interconversion of the low-pH and high-pH conformations of cathepsin D is a gradual transition process over a broad pH range (*33*). On the basis of these facts, we consider two possible modes of inhibitory interaction between synthetic peptide [1−10] derived from the mature N-terminus and cathepsin D. Depending on the pH, it can bind either to the active site or to the low-pH binding site. In both cases, the peptide competes with the intact mature N-terminus and influences its transition relocation. Thus, peptide [1−10] can directly or indirectly reduce the accessibility of the active site for substrate. According to the X-ray structures, both binding modes of the mature N-terminus involve critical interactions through Lys8 and Tyr10. This is in line with our finding that substitution of these residues in peptide [1−10] abolishes its inhibition. The noncompetitive inhibition measured for peptide [1−10] at pH 5.1 suggests that at this pH, the inhibition through the indirect pathway is prevalent since the same type of inhibition was also identified for the N-terminus of propeptide.

We conclude that three autonomous inhibitory domains are involved in the specific inactivation of the enzyme core of cathepsin D on the zymogen level and mature enzyme level. The analysis of these domains provided new insight into the mechanism of the activation process and its potential modulation through propeptide mutation and interaction of propeptide with glycosaminoglycans. The discovered inhibitory structures represent promising leads for designing a new class of propeptide-derived inhibitors of cathepsin D.

ACKNOWLEDGMENT

We thank I. Pražáková for technical assistance, M. Blechová for peptide synthesis, P. Maloň and H. Dlouhá for circular dichroism spectroscopy, J. Cvačka and J. Kohoutová for mass spectrometry, J. Zbrožek and V. Himrová for amino acid analysis, and M. Baudyš and H. Hoffman for critical reading of the manuscript.

REFERENCES

1. Dunn, B. M. (2002) Structure and mechanism of the pepsin-like family of aspartic peptidases, *Chem. Rev. 102*, 4431−4458.
2. Dunn, B. M., Deyrup, C., Moesching, W. G., Gilbert, W. A., Nolan, R. J., and Trach, M. L. (1978) Inhibition of pepsin by zymogen activation fragments. Spectrum of peptides released from pepsinogen NH_2 terminus and solid phase synthesis of two inhibitory peptide sequences, *J. Biol. Chem. 253*, 7269−7275.
3. Evin, G., Devin, J., Castro, B., Menard, J., and Corvol, P. (1984) Synthesis of peptides related to the prosegment of mouse submaxillary gland renin precursor: An approach to renin inhibitors, *Proc. Natl. Acad. Sci. U.S.A. 81*, 48−52.
4. Cumin, F., Evin, G., Fehrentz, J. A., Seyer, R., Castro, B., Menard, J., and Corvol, P. (1985) Inhibition of human renin by synthetic peptides derived from its prosegment, *J. Biol. Chem. 260*, 9154−9157.
5. Fusek, M., Mares, M., Vagner, J., Voburka, Z., and Baudys, M. (1991) Inhibition of aspartic proteinases by propart peptides of human procathepsin D and chicken pepsinogen, *FEBS Lett. 287*, 160−162.
6. Koelsch, G., Mares, M., Metcalf, P., and Fusek, M. (1994) Multiple functions of pro-parts of aspartic proteinase zymogens, *FEBS Lett. 343*, 6−10.
7. Wittlin, S., Rosel, J., and Stover, D. R. (1998) One-step purification of cathepsin D by affinity chromatography using immobilized propeptide sequences, *Eur. J. Biochem. 252*, 530−536.
8. Saftig, P., Hetman, M., Schmahl, W., Weber, K., Heine, L., Mossmann, H., Koster, A., Hess, B., Evers, M., von Figura, K., and Peters, C. (1995) Mice deficient for the lysosomal proteinase cathepsin D exhibit progressive atrophy of the intestinal mucosa and profound destruction of lymphoid cells, *EMBO J. 14*, 3599−3608.
9. Steinfeld, R., Reinhardt, K., Schreiber, K., Hillebrand, M., Kraetzner, R., Bruck, W., Saftig, P., and Gartner, J. (2006) Cathepsin D deficiency is associated with a human neurodegenerative disorder, *Am. J. Hum. Genet. 78*, 988−998.
10. Tyynela, J., Sohar, I., Sleat, D. E., Gin, R. M., Donnelly, R. J., Baumann, M., Haltia, M., and Lobel, P. (2000) A mutation in the ovine cathepsin D gene causes a congenital lysosomal storage disease with profound neurodegeneration, *EMBO J. 19*, 2786−2792.
11. Papassotiropoulos, A., Bagli, M., Kurz, A., Kornhuber, J., Forstl, H., Maier, W., Pauls, J., Lautenschlager, N., and Heun, R. (2000) A genetic variation of cathepsin D is a major risk factor for Alzheimer's disease, *Ann. Neurol. 47*, 399−403.
12. Davidson, Y., Gibbons, L., Pritchard, A., Hardicre, J., Wren, J., Tian, J., Shi, J., Stopford, C., Julien, C., Thompson, J., Payton, A., Thaker, U., Hayes, A. J., Iwatsubo, T., Pickering-Brown, S. M., Pendleton, N., Horan, M. A., Burns, A., Purandare, N., Lendon, C. L., Neary, D., Snowden, J. S., and Mann, D. M. (2006) Genetic associations between cathepsin D exon 2 C → T polymorphism and Alzheimer's disease, and pathological correlations with genotype, *J. Neurol., Neurosurg. Psychiatry 77*, 515−517.
13. Bossard, N., Descotes, F., Bremond, A. G., Bobin, Y., De Saint Hilaire, P., Golfier, F., Awada, A., Mathevet, P. M., Berrerd, L., Barbier, Y., and Esteve, J. (2003) Keeping data continuous when

analyzing the prognostic impact of a tumor marker: An example with cathepsin D in breast cancer, *Breast Cancer Res. Treat. 82*, 47−59.
14. Vetvicka, V., Vetvickova, J., Hilgert, I., Voburka, Z., and Fusek, M. (1997) Analysis of the interaction of procathepsin D activation peptide with breast cancer cells, *Int. J. Cancer 73*, 403−409.
15. Glondu, M., Coopman, P., Laurent-Matha, V., Garcia, M., Rochefort, H., and Liaudet-Coopman, E. (2001) A mutated cathepsin-D devoid of its catalytic activity stimulates the growth of cancer cells, *Oncogene 20*, 6920−6929.
16. Beaujouin, M., Baghdiguian, S., Glondu-Lassis, M., Berchem, G., and Liaudet-Coopman, E. (2006) Overexpression of both catalytically active and inactive cathepsin D by cancer cells enhances apoptosis-dependent chemo-sensitivity, *Oncogene 23*, 1967−1973.
17. Metcalf, P., and Fusek, M. (1993) Two crystal structures for cathepsin D: The lysosomal targeting signal and active site, *EMBO J. 12*, 1293−1302.
18. Baldwin, E. T., Bhat, T. N., Gulnik, S., Hosur, M. V., Sowder, R. C., II, Cachau, R. E., Collins, J., Silva, A. M., and Erickson, J. W. (1993) Crystal structures of native and inhibited forms of human cathepsin D: Implications for lysosomal targeting and drug design, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6796−6800.
19. Lee, A. Y., Gulnik, S. V., and Erickson, J. W. (1998) Conformational switching in an aspartic proteinase, *Nat. Struct. Biol. 5*, 866−871.
20. Khan, A. R., Cherney, M. M., Tarasova, N. I., and James, M. N. (1997) Structural characterization of activation 'intermediate 2' on the pathway to human gastricsin, *Nat. Struct. Biol. 4*, 1010−1015.
21. Ostermann, N., Gerhartz, B., Worpenberg, S., Trappe, J., and Eder, J. (2004) Crystal structure of an activation intermediate of cathepsin E, *J. Mol. Biol. 342*, 889−899.
22. Hartsuck, J. A., Koelsch, G., and Remington, S. J. (1992) The high-resolution crystal structure of porcine pepsinogen, *Proteins 13*, 1−25.
23. Moore, S. A., Sielecki, A. R., Chernaia, M. M., Tarasova, N. I., and James, M. N. (1995) Crystal and molecular structures of human progastricsin at 1.62 Å resolution, *J. Mol. Biol. 247*, 466−485.
24. Fusek, M., Baudys, M., and Metcalf, P. (1992) Purification and crystallization of human cathepsin D, *J. Mol. Biol. 226*, 555−557.
25. Cheng, Y., and Prusoff, W. H. (1973) Relationship between the inhibition constant (K_I) and the concentration of inhibitor which causes 50 percent inhibition (I_{50}) of an enzymatic reaction, *Biochem. Pharmacol. 22*, 3099−3108.
26. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A 2nd Generation Force Field for the Simulation of Proteins and Nucleic Acids, *J. Am. Chem. Soc. 117*, 5179−5197.
27. Lakowicz, J. R. (1999) *Principles of fluorescence spectroscopy*, Plenum, New York.
28. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera: A Visualization System for Exploratory Research and Analysis, *J. Comput. Chem. 25*, 1605−1612.
29. Richo, G. R., and Conner, G. E. (1994) Structural requirements of procathepsin D activation and maturation, *J. Biol. Chem. 269*, 14806−14812.
30. Wittlin, S., Rosel, J., Hofmann, F., and Stover, D. R. (1999) Mechanisms and kinetics of procathepsin D activation, *Eur. J. Biochem. 265*, 384−393.
31. Larsen, L. B., Boisen, A., and Petersen, T. E. (1993) Procathepsin D cannot autoactivate to cathepsin D at acid pH, *FEBS Lett. 319*, 54−58.
32. Lazo, N. D., and Downing, D. T. (1997) Circular dichroism of model peptides emulating the amphipathic α-helical regions of intermediate filaments, *Biochemistry 36*, 2559−2565.
33. Goldfarb, N. E., Lam, M. T., Bose, A. K., Patel, A. M., Duckworth, A. J., and Dunn, B. M. (2005) Electrostatic switches that mediate the pH-dependent conformational change of "short" recombinant human pseudocathepsin D, *Biochemistry 44*, 15725−15733.
34. Touitou, I., Capony, F., Brouillet, J. P., and Rochefort, H. (1994) Missense polymorphism (C/T224) in the human cathepsin D profragment determined by polymerase chain reaction-single strand conformational polymorphism analysis and possible consequences in cancer cells, *Eur. J. Cancer 30*, 390−394.
35. Yasuda, Y., Li, Z., Greenbaum, D., Bogyo, M., Weber, E., and Bromme, D. (2004) Cathepsin V, a novel and potent elastolytic activity expressed in activated macrophages, *J. Biol. Chem. 279*, 36761−36770.
36. Almeida, P. C., Nantes, I. L., Chagas, J. R., Rizzi, C. C., Faljoni-Alario, A., Carmona, E., Juliano, L., Nader, H. B., and Tersariol, I. L. (2001) Cathepsin B activity regulation. Heparin-like glycosaminogylcans protect human cathepsin B from alkaline pH-induced inactivation, *J. Biol. Chem. 276*, 944−951.
37. Rozman, J., Stojan, J., Kuhelj, R., Turk, V., and Turk, B. (1999) Autocatalytic processing of recombinant human procathepsin B is a bimolecular process, *FEBS Lett. 459*, 358−362.
38. Vasiljeva, O., Dolinar, M., Turk, V., and Turk, B. (2003) Recombinant human cathepsin H lacking the mini chain is an endopeptidase, *Biochemistry 42*, 13522−13528.

BI0614986