Hydrogen-bond distance restraints were determined from the temperature coefficients of the backbone amide NH resonances, measured from a series of HSQC spectra recorded between 278 and 298 K. ^{15}N T_1, T_2 and $^1H/^{15}N$ NOE parameters were determined using gradient sensitivity-enhanced minimal water saturation pulse sequences[24]. The self-diffusion coefficient for MKLysEPO was measured as described in ref. 10. Structures were calculated with a dynamical simulated annealing protocol[25] using the X-PLOR program[26]. The structure calculations employed 1,741 proton-proton distance restraints, 124 hydrogen-bond restraints derived for 62 amide NHs, and 85 ϕ angle restraints. NOE-derived restraints were conservatively categorized as strong, medium and weak, with upper bounds of 2.7, 3.3 and 5.0 Å respectively, on the basis of observed NOE intensities.

Coordinates. Coordinates for the NMR structure of MKLysEPO have been deposited in the Brookhaven Protein Databank (accession code 1buy).

Acknowledgments
The authors gratefully acknowledge T. Boone for providing the original LysEPO clone, G. Stearns and G. Rogers (AMGEN) for help in production of labeled protein and L.E. Kay (University of Toronto) for initial NMR data collection facilities. We also thank N.A. Farrow (University of Toronto) and J. Cavanagh (New York State Dept. Health) for discussions on NMR relaxation data analysis, and S. Elliott and T. Osslund (AMGEN) for many useful discussions and critical reading of the manuscript.

Janet C. Cheetham, Duncan M. Smith, Kenneth H. Aoki, Janice L. Stevenson, Thomas J. Hoeffel, Rashid S. Syed, Joan Egrie and Timothy S. Harvey.

Amgen Inc, One Amgen Center Drive, Thousand Oaks, California 91320, USA.

Correspondence should be addressed to J.C.C. *email: cheetham@amgen.com*

Received 4 August, 1998; accepted 8 August, 1998.

1. Graber, S.E. & Krantz, S.B. *Annu. Rev. Med.* **29**, 51 (1978).
2. Youssoufian, H., Longmore, G., Neumann, D., Yoshimura, A. & Lodish, H.F. *Blood* **81**, 2223–2223 (1993).
3. Markham, A. & Bryson, H.M. *Drugs* **49**, 232–254 (1995).
4. Lin, F.K. *et al. Proc. Natl. Acad. Sci. USA* **82**, 7580–7584 (1985).
5. Narhi, L.O. *et al. J. Biol. Chem.* **266**, 23022–23026 (1991).
6. Altieri, A.S., Hinton, D.P. & Byrd, R.A. *J. Am. Chem. Soc.* **117**, 7566–7567 (1995).
7. Sprang, S.R. & Bazan, J.F. *Curr. Opin. Struct. Biol.* **3**, 815–827 (1993).
8. Harris, N.L., Presnell, S.R. & Cohen, F.E. *J. Mol. Biol.* **236**, 1356–1368 (1994).
9. Wishart, D.S. & Sykes, B.D. *J. Biomol. NMR* **4**, 171–180 (1994).
10. Elliott, S. *et al. Blood* **87**, 2702–2713 (1996).
11. Wen, D., Boissel, J.P., Showers, M., Ruch, B.C. & Bunn, H.F. *J. Biol. Chem.* **269**, 22839–22846 (1994).
12. Zhang, F. *et al. Nature* **387**, 206–209 (1997).
13. Elliott, S., Lorenzini, T., Chang, D., Barzilay, J. & Delorme, E. *Blood* **89**, 493–502 (1997).
14. Bittorf, T., Jaster, R. & Brock, J. *FEBS Lett.*. **336**, 133–136 (1993).
15. Matthews, D.J., Topping, R.S., Cass, R.T. & Giebel, L.B. *Proc. Natl. Acad. Sci. USA* **93**, 9471–9476 (1996).
16. Syed, R.S. *et al. Nature*, **in the press** (1998).
17. de, V.-A.M., Ultsch, M. & Kossiakoff, A.A. *Science* **255**, 306–312 (1992).
18. Bazan, J.F. *ImmunoL. Today* **11**, 350–354 (1990).
19. Robinson, R.C. *et al. Cell* **77**, 1101–1116 (1994).
20. Xu, G.Y. *et al. J. Mol. Biol.* **268**, 468–481 (1997).
21. Kay, L.E. *Prog. Biophys. Mol. Biol.* **63**, 277–299 (1995).
22. Zink, T. *et al. Biochemistry* **33**, 8453–8463 (1994).
23. Kuboniwa, H., Grzesiek, S., Delaglio, F. & Bax, A. *J. Biomol. NMR* **4**, 871–878 (1994).
24. Farrow, N.A. *et al. Biochemistry* **33**, 5984–6003 (1994).
25. Bagby, S., Harvey, T.S., Eagle, S.G., Inouye, S. & Ikura, M. *Structure* **2**, 107–122 (1994).
26. Brunger, A.T. *X-PLOR 3.1 Manual* (Yale University Press, New Haven, Connecticut; 1992).
27. Esnouf, R.M. *J. Mol. Graphics* **15**, 133–138 (1997).
28. Ghose, A. & Crippen, G. *J. Comp. Chem.* **7**, 565–577 (1986).
29. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. *J. Appl. Crystallogr.* **26**, 283–291 (1993).

Conformational switching in an aspartic proteinase

The crystal structure of a catalytically inactive form of cathepsin D ($CatD_{hi}$) has been obtained at pH 7.5. The N-terminal strand relocates by 30 Å from its position in the interdomain β-sheet and inserts into the active site cleft, effectively blocking substrate access. $CatD_{hi}$ has a five-stranded interdomain β-sheet and resembles Intermediate 3 , a hypothetical structure proposed to be transiently formed during proteolytic activation of the proenzyme precursor. Interconversion between active and inactive forms of CatD is reversible and may be regulated by an ionizable switch involving the carboxylate side chains of Glu 5, Glu 180, and Asp 187. Our findings provide a structural basis for the pH-dependent regulation of aspartic proteinase activity and suggest a novel mechanism for pH-dependent modulation of substrate specificity.

Aspartic proteinases represent a large family of enzymes that catalyze peptide bond hydrolysis through an acid-base mechanism mediated by two catalytic aspartic acid residues[1,2]. They are synthesized at neutral pH as inactive zymogens that are inhibited by binding of their propeptide segments in the active site. The proenzymes can be autoactivated to their mature forms upon acidification, which leads to a concerted series of conformational changes concomitant with both intra- and inter-molecular processing events[1]. Although the mature enzymes share the same catalytic machinery and overall characteristic fold, they display wide variations in pH-dependency for stability and substrate catalysis, the basis of which has been a subject of some debate.

Cathepsin D (CatD) is a lysosomal aspartic proteinase involved in the turnover of cellular proteins as well as in the selective processing of MHC II antigens[3], hormones and growth factors[4]. It has been shown that CatD can cleave β-amyloid precursor protein to release an amyloidogenic peptide found in the amyloid plaques of Alzheimers disease patients[5,6]. Over-expression of CatD has been implicated in progressive breast cancer[7]. Recent gene knockout studies have demonstrated that CatD-deficient mice die prematurely within four weeks from massive physiological abnormalities[8]. Recognition of the potential roles of CatD in human diseases has led several groups to design inhibitors[9,10] that will help to delineate the role of this enzyme in normal and disease processes.

The crystal structures of native and a pepstatin-inhibited form of human CatD have been solved[11] to aid in the structure-based design of CatD inhibitors[9,10]. Crystal structures of bovine and human CatD from different sources are very similar[12]. All of these structures were obtained from crystals grown at acidic pH. There has been a great deal of controversy over whether CatD can function in the extracellular environment due to its low pH optimum. In this paper, we describe the crystal structure of a novel form of CatD obtained at neutral pH, and discuss the potential biological significance of this structure for enzyme activation and substrate specificity.

pH-induced conformation changes in cathepsin D

Low pH crystal structures for both native and pepstatin-inhib-

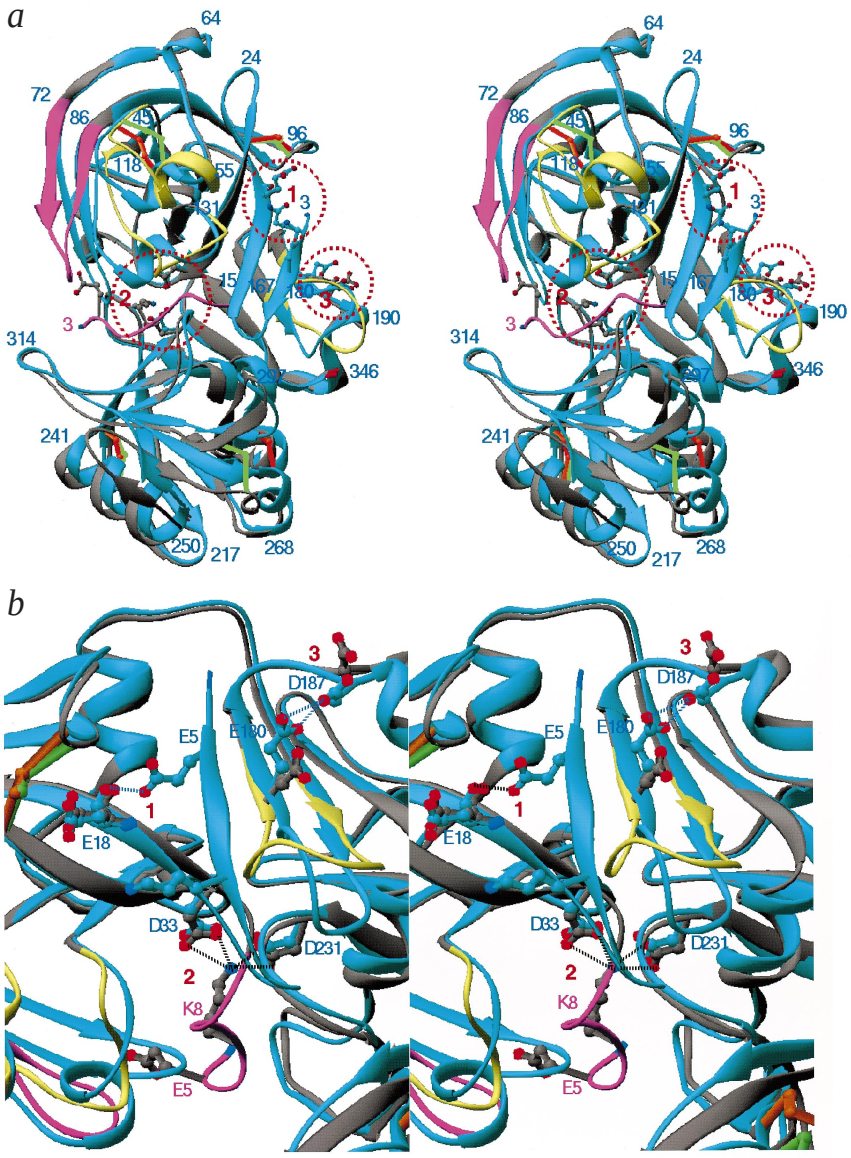

Fig. 1 Stereoview of ribbon superposition of the high and low pH forms of CatD. ***a***, The superposed crystal structures for the low pH form, $CatD_{lo}$ (cyan) and the high pH form, $CatD_{hi}$ (gray) with the locations of putative ionizable molecular switches shown in the circles (numbered 1–3). The four disulfide bonds are also shown for $CatD_{lo}$ (orange) and $CatD_{hi}$ (green). The structural alignment was done using the superposition method of Rossman and Argos[33]. All Cα atoms pairs with a distance of <0.52 Å were chosen, resulting in 164 out of 338 pairs (residues 17–31, 33–44, 57–64, 66–71, 86–95, 108–118, 132–144, 146–160, 163–167, 181–185, 192–211, 213–217, 227–233, 271–284, and 328–345). The ordered N-terminal segment (residues 3–13) and the flap of $CatD_{hi}$ (residues 72–86) are colored in magenta; the two N-termini are indicated by residue number 3. Other segments that undergo localized conformational changes unique to the $CatD_{hi}$ structure are colored in yellow, and side chains for residues in the molecular switch are drawn in ball-and-stick. ***b***, Close-up stereoview of the putative inozable switches shown in the circles of Fig. 1*a*. Figures were produced with the program RIBBONS[34].

ited human CatD are nearly identical and exhibit a typical aspartic proteinase fold[11,12]. We obtained a novel crystal form of the native enzyme at pH 7.5, at which the enzyme is catalytically inactive. The crystal structure of this high pH form, designated $CatD_{hi}$, was determined to 2.5 Å resolution by molecular replacement. Structural comparison of $CatD_{hi}$ and the pH 5.1 form of the enzyme, designated $CatD_{lo}$, reveals a number of striking conformational rearrangements (Fig. 1*a*).

The N-terminal segment (residues 3–7), comprising the first strand of the interdomain β-sheet in $CatD_{lo}$, is rearranged to a flexible coil and inserted into the active site cleft in $CatD_{hi}$. Thus, the interdomain β-sheet in $CatD_{hi}$ consists of only five-strands instead of the normal complement of six-strands present in both mature and proenzyme forms of aspartic proteinases. The relocation of the N-terminus can be achieved by rotation of nearly 180° about a hinge located at Ala 13 with associated disruption of the β-hairpin turn, resulting in a maximum displacement of up to 30 Å. Residues 3–16 of $CatD_{hi}$ have well-defined continuous electron density (Fig. 2*a*) and span the substrate binding cleft, thus effectively preventing access by substrate or inhibitor. The N-terminal sequence of the mature enzyme was confirmed by Edman degradation, implying that the first two residues were disordered in the structure of $CatD_{hi}$. The charged Nζ atom of Lys 8 from the N-terminal strand is located in a position normally occupied by either a catalytic water molecule or the OH group of a peptidomimetic inhibitor, and it forms a salt bridge with the catalytic carboxylates, Asp 33 and Asp 231, which maintain a coplanar geometry (Fig. 1*b*). The backbone of the N-terminal strand follows a different course through the active site cleft when compared with that of a peptidomimetic inhibitor[11]. As a result, the S1 specificity pocket is altered by the insertion of Tyr 10 whose hydroxyl group makes a hydrogen bond with the side chain of Asp 33. In addition, residues Pro 4 and Glu 5 displace the β-hairpin structure, termed the 'flap' (residues 72–86), by up to 8.0 Å from its position in $CatD_{lo}$ (Fig. 1*a*). The electron density for the flap is clearly defined in the final omit ($|F_o| - |F_c|$) electron density map (Fig. 2*b*).

The flap is one of the most mobile elements in aspartic proteinase structures, as evidenced by its relatively high tempera-

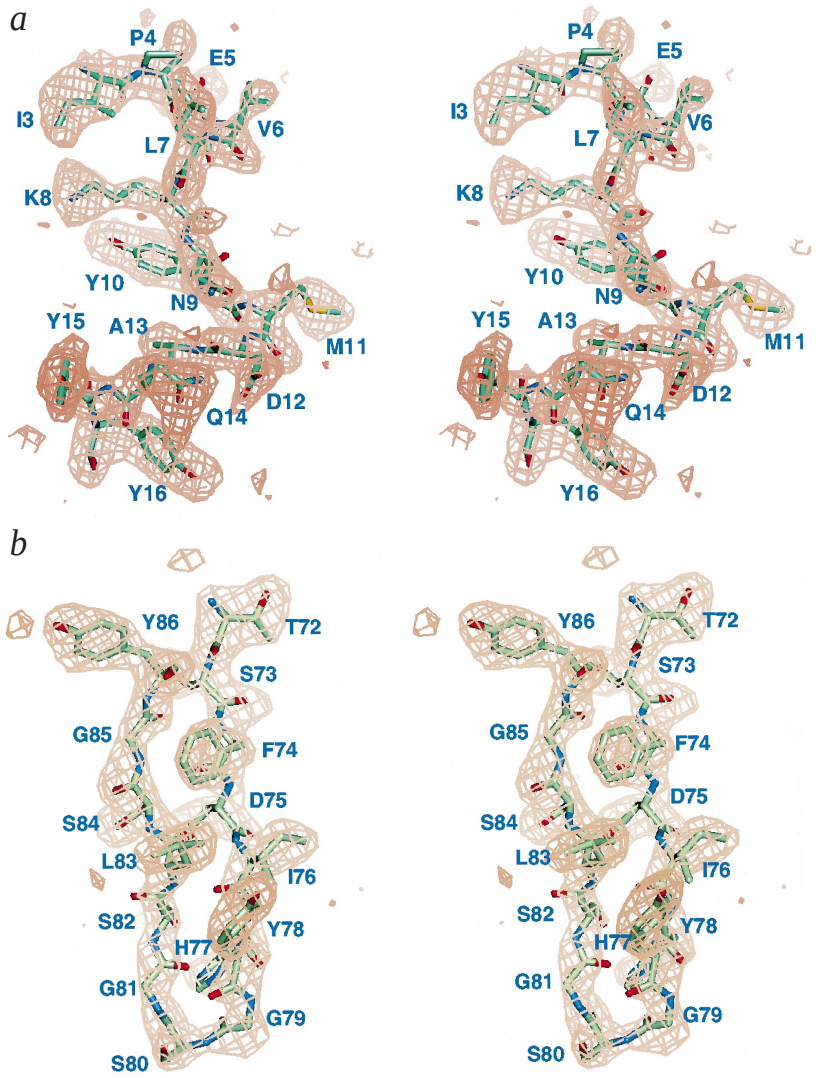

Fig. 2 Stereoview of the calculated electron density for omit F_o - F_c electron density maps corresponding to ***a***, the N-terminal peptide residues, Ile 3–Tyr 16; and ***b***, the flap residues, Thr 72–Phe 86. The N-terminal peptides and flap were omitted from simulated annealing refinement and from the map calculations. Maps were contoured at 2.5σ (*a*) and 2.8σ (*b*). The figures were prepared with the program QUANTA97 (MSI/Biosym, Inc).

ture factors and minor differences between the tips of the flaps in native and inhibitor bound forms of the enzymes[1]. The magnitude of the flap movement observed in $CatD_{hi}$, however, is unprecedented and reminiscent of the flap movements seen in retroviral proteinases[13]. In addition, the relocation of the N-terminal strand is accompanied by a conformational change in the neighboring segment, Ala 118–Phe 131, that exhibits a maximum displacement of 5.1 Å from its location in $CatD_{lo}$ and forms a portion of the hydrophobic binding pocket that encompasses the side chain of Tyr 10. The shift in location of Ala 118–Phe 131 is also correlated with a conformational change in an adjacent segment, His 45–His 56, with which it makes extensive van der Waals interactions. The latter segment contains a surface loop formed by the disulfide linkage between Cys 46 and Cys 53, that undergoes a large backbone conformational change in which Leu 49 moves up to 6.2 Å relative to its position in $CatD_{lo}$. Finally, the surface loop formed by residues 169–181 shifts by up to 4.0 Å to allow this segment to partially occupy the region vacated by the N-terminus.

Besides the localized conformational changes between $CatD_{lo}$ and $CatD_{hi}$, comparison of the two structures also reveals a major shift in the relative orientations of their N- and C-domains, that can be described by a rigid-body rotation of 7.2° and a translation of 1.5 Å, an unusually large domain movement for an aspartic proteinase[14,15]. Previously reported structures of both native and inhibited forms of mature CatD illustrate only minor domain displacements for this enzyme[16]. The $CatD_{hi}$ structure, however, suggests that CatD can display high degree of segmental flexibility under certain conditions.

Ionizable switch

Since the structures of $CatD_{hi}$ and $CatD_{lo}$ were obtained at different pHs, we searched for the existence of potential pH-dependent ionizable switches (Fig. 1*a*,*b*). One obvious candidate is the interacting pair of catalytic aspartic acid residues, Asp 33 and Asp 231, that may both be deprotonated at pH 7.5, resulting in repulsion. In the structure of $CatD_{hi}$, however, the insertion of Lys 8 into the active site effectively neutralizes the additional negative charge. Other possible switches in $CatD_{lo}$ may involve interactions between the two carboxylates of Glu 180 and Asp 187, and between the carboxylate of Glu 5 and the carbonyl oxygen of Glu 18. Both of these sets of interactions would become unfavorable at high pH and are not present in the $CatD_{hi}$ structure (Fig. 1*b*). The carboxylate side chains of Glu 180 and Asp 187, which are within hydrogen bonding distance in $CatD_{lo}$, are over 8 Å apart in $CatD_{hi}$. The rearrangement of the N-terminal strand also eliminates the potential repulsive electrostatic interaction between Glu 5 and Glu 18. It is noteworthy that Glu 180 and Asp 187, as well as Glu 5, are conserved among CatD sequences from different species, consistent with an important role for these residues.

Structural similarity to the proenzyme

The insertion of the N-terminal strand into the active site in $CatD_{hi}$ is reminiscent of the mechanism by which the zymogen precursors of aspartic proteinases are maintained in an inactive state[17–19]. Activation of zymogen precursors of enzymes to their mature forms upon acidication is a complex multi-step process that requires multiple cleavage events and major conformational rearrangements of the protein. During the activation, the mature N-terminus of the enzyme relocates by ~40 Å and displaces the N-terminus of the propeptide from the outermost sixth strand of the interdomain β-sheet by a kind of strand exchange reaction[17]. Structural comparison among the active site regions of $CatD_{hi}$, pepsinogen (the precursor of

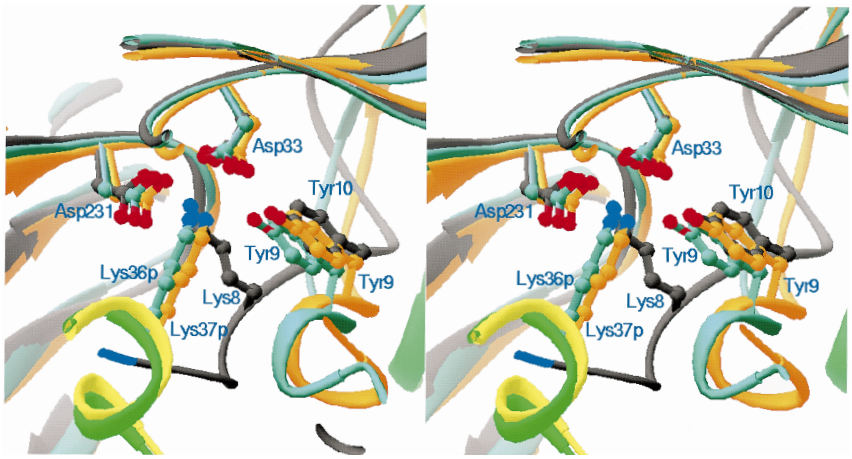

Fig. 3 Stereoview of structural alignments of the active sites for pepsinogen (PDBBANK: 2PSG)[17], progastricsin (PDBBANK: 1HTR)[19], and $CatD_{hi}$. The pepsinogen and progastricsin are colored in light blue and orange, except for their propeptide segments (green for pepsinogen, yellow for progastricsin), while $CatD_{hi}$ is in gray. The ε-NH_3 group of Lys 8 in $CatD_{hi}$ occupies the same position as that of Lys 36p in the propeptide segments of pepsinogen and progastricsin (Lys 37p). Tyr 10 in $CatD_{hi}$ occupies the same binding pocket as Tyr 9 in pepsinogen and progastricsin, and all forms stabilizing hydrogen bonds with Asp 33 (CatD numbering) in the three structure.

pepsin), and progastricsin (the precursor of gastricsin), reveals that the two pairs of catalytic carboxylates lie in approximately the same plane (Fig. 3). Lys 36p from the N-terminal peptide segment of pepsonigen (Lys 37p in progastrincsin) makes ionic interactions with the catalytic aspartates, blocking the productive binding of substrate. The charged Nζ atom of Lys 8 in $CatD_{hi}$, despite coming from a different direction and portion of the structure, occupies the same position as that of Lys 36p in pepsinogen (Lys 37p in progastricsin). The side chain of Tyr 9 in pepsinogen and progastricsin nearly superimposes with that of Tyr 10 in $CatD_{hi}$. The hydroxyl groups of these three tyrosines form a hydrogen bond with their corresponding catalytic aspartic acid residues: Asp 33 in $CatD_{hi}$ and Asp 32 in pepsinogen and progastricsin. Superposition of pepsinogen with pepsin reveals a large shift in the position of Tyr 9 as a result of the activation process[17,18]. We propose that both Lys 8 and Tyr 10 play critical roles in stabilizing the high pH, inactive form of mature CatD in the manner similar to that of their counterparts, Lys 36p and Tyr 9 (Lys 37p and Tyr 9), that stabilize the inactive zymogen in pepsinogen (progastricsin).

Recently, the crystal structure of 'intermediate 2' in the activation of human progastricsin has been reported[20]. Intermediate 2 results from the cleavage at the primary and secondary propeptide processing site but has the N-terminal propeptide segment non-covalently associated as part of the interdomain β-sheet secondary structural element. In this respect, intermediate 2 represents a novel inhibited form of the mature enzyme in which the propeptide has displaced the N-terminus but lacks the active site neutralizing Lys 37p residue. Propeptide segments have been reported to be inhibitors for several aspartic proteinases including pespin, renin and cathepsin D[21–24]. The N-terminal portion of the propeptide can bind to mature cathepsin D[25]. In certain respects, the $CatD_{hi}$ structure resembles the structure of 'intermediate 3' which has been proposed to be a transient intermediate during zymogen activation[25]. Intermediate 3 is predicted to contain a five-stranded interdomain β-sheet, resulting from autolysis and release of the nascent propeptide segments. In the case of $CatD_{hi}$, this intermediate form can apparently be stabilized by the Lys-X-Tyr motif in the N-terminal strand which functionally replaces these key stabilizing residues in the zymogen precursors. The existence of a five-stranded form of an aspartic

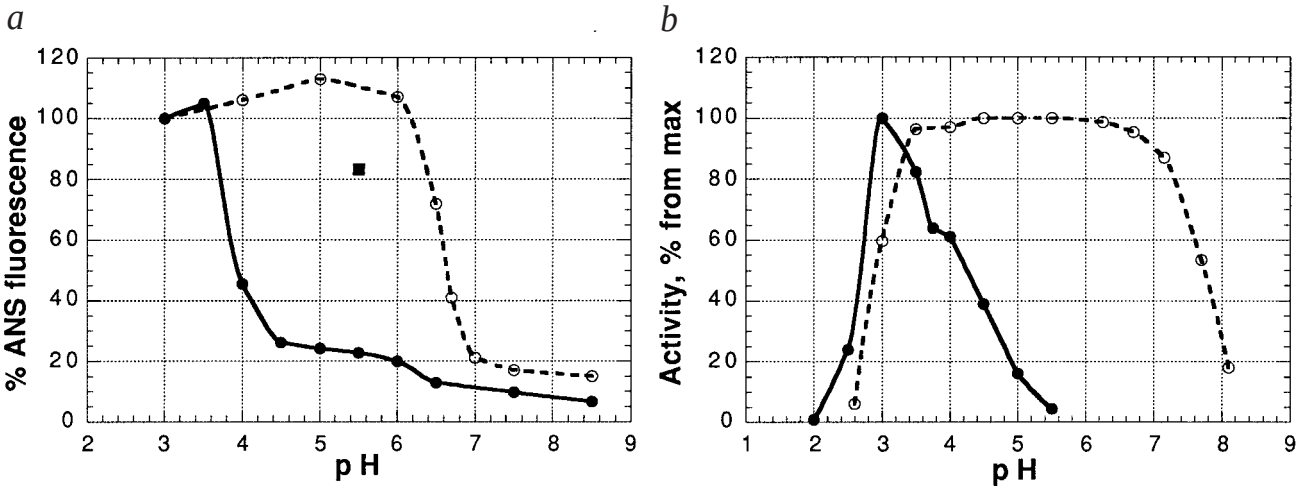

Fig. 4 pH dependence of CatD conformation and activity in solution. ***a***, Relative ANS fluorescence upon binding to native CatD (solid line) and to a CatD–pepstatin A complex (dashed line) as a function of pH. The solid square indicates relative fluorescence obtained upon ANS binding to CatD in the presence of 2.4 M ammonium sulfate. The change in ANS fluorescence was expressed as percent relative to that observed at pH 3.0. ***b***, Enzyme activity (solid line) and enzyme stability (dash line) of CatD as a function of pH. Activity was determined as described in the Methods section and expressed as percent of the maximal activity obtained at pH 3.0 (solid circles). Enzyme stability was measured by monitoring hemoglobin degradation at pH 3.0 after preincubating the enzyme for 100 min at 37 ºC at the indicated pH (open circles). Maximum activity (100%) was equal to initial activity before the pre-incubation.

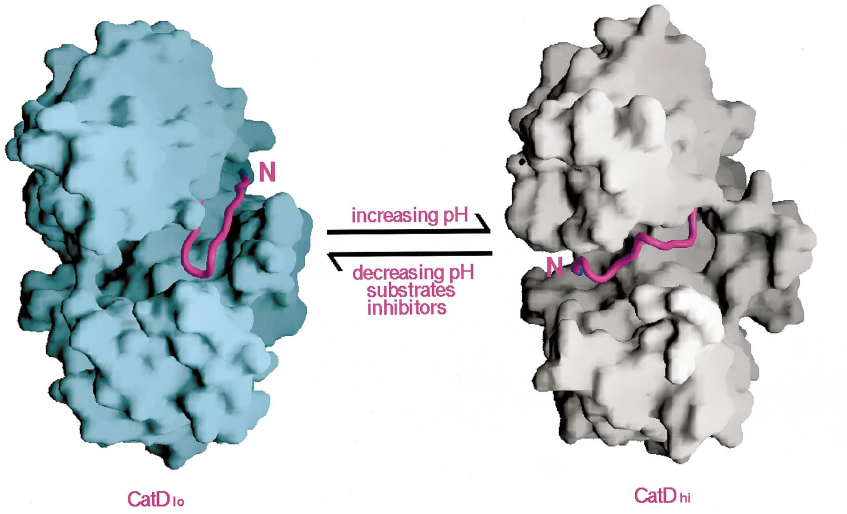

Fig. 5 Molecular surfaces illustrating the pH and substrate dependence of the conformational equilibrium between the $CatD_{lo}$ and $CatD_{hi}$. The N-terminal segment (residues 3–17) is colored in magenta. Images were created using the program GRASP[35] and coordinates (PDBBANK: 1LYB)[11] for $CatD_{lo}$ and those of the current described structure of $CatD_{hi}$ in the same orientation.

proteinase provides a satisfying glimpse of a requisite but unseen intermediate in the strand exchange reaction that occurs during zymogen activation.

Spectroscopic studies of reversibility

The fluorescent probe, 8-anilino-1-naphthalenesulfonate (1,8-ANS), is often used to monitor conformational changes of proteins in solution[26]. CatD exhibits a strongly pH-dependent binding profile for ANS (Fig. 4*a*). The maximum binding for native enzyme is obtained at pH 3.0–3.5, near the pH optimum of the enzyme (Fig. 4*b*). The reduction in ANS binding observed at high pH may reflect the conformational change of $CatD_{lo}$ to $CatD_{hi}$ and account for the loss in enzyme activity at higher pH. The more rapid decrease in ANS binding *versus* enzyme activity with increasing pH may be due to substrate stabilization of $CatD_{lo}$. In this regard, pepstatin A, a potent inhibitor of CatD, increases the pH at which the transition of $CatD_{lo}$ to $CatD_{hi}$ occurs (Fig. 4*a*). The pK_a for ANS dissociation for the CatD–pepstatin A complex is around pH 6.5 and is probably due to ionization of the second active site carboxylate group which is presumed to have an anomalously high pK_a. If this is true, how can one explain the fact that the crystal structure of native, active $CatD_{lo}$ is obtained from crystals grown at pH 5.1 which should favor the inactive $CatD_{hi}$ species according to our data (Fig. 4*b*)? One possible explanation is that the presence of 2.4 M ammonium sulfate, used in the crystallization of CatD at pH 5.1 in the absence of bound ligand, actually stabilizes the $CatD_{lo}$ form, and this is supported by the ANS binding data (Fig. 4*a*, black square).

An autoregulatory mechanism for proteinase activity

The catalytic mechanism of aspartic proteinases has been the subject of numerous enzymatic and structural studies[1]. While it is widely accepted that the mechanism depends on the protonation state of the two catalytic aspartates, it is not as well understood how each enzyme exhibits its own unique pH profile for optimal proteolytic activity. Our crystallographic and biochemical results with CatD suggest a novel structural mechanism for the pH-dependent regulation of aspartic proteinase activity. Electrostatic repulsion in $CatD_{lo}$, induced by an increase in pH, results in a concerted set of conformational changes including the relocation of the N-terminus into the active site. The trigger for the structural rearrangement may be an ionizable switch formed by the interaction between Glu 180 and Asp 187, and possibly Glu 5 and the backbone of Glu 18. The N-terminal segment in $CatD_{hi}$ is inserted into the active site, where both catalytic aspartates are likely deprotonated. Ionic interactions with the positively charged Nζ atom of Lys 8 along with the hydrogen bond formed with the hydroxyl moiety of Tyr 10 stabilize the ionized carboxylates. Thus, Lys 8 and Tyr 10 of $CatD_{hi}$ represent an inhibitory motif that blocks the substrate binding site. Reactivation of $CatD_{hi}$ upon acidification could occur due to the protonation of one of the two catalytic aspartates, thus weakening their electrostatic interactions with Lys 8. The released N-terminal peptide reforms the sixth strand of the antiparallel β-sheet and restores the accessibility of the active site to substrate. This mechanism is consistent with our observations that CatD exhibits fully reversible catalytic activity up to pH 7.0 with 50% recoverable activity at pH 7.5 (Fig. 4*b*).

Potential biological roles of the autoregulation

What is the biological significance of the pH-dependent autoregulation of CatD? Experiments with CatD gene knockout mice[8] suggest that CatD may not be critical for protein catabolism, which is thought to be one its major functions. Instead, it has been proposed that CatD may be important for more specific and restricted proteolytic processing events such as the activation of proenzymes, prohormones and growth factors[4,8]. If so, there should be a mechanism for the regulation of its enzymatic activity. CatD is markedly unstable at pH 3.0 where optimal enzyme activity is observed, presumably a reflection of autolysis. At pH 4–6, the enzyme is quite stable, even though it still possesses significant enzyme activity over part of this range. CatD is found in extremely high concentrations, up to 20 mg ml^{-1}, in the lysosome where the pH has been reported to be in the range 4–6[27]. Recent studies demonstrate that elevated levels of mature CatD can be found in extralysosomal environments, such as cytosol and blood, either after exhaustive physical exercise[28] or in certain pathological situations[4]. The pH in the bloodstream is normally regulated to pH 7.4 ± 0.03, and the pH of the cytosol is also maintained within very narrow limits near neutrality, pH 7.2–7.4. We propose that the active, low pH and the inactive, high pH forms of CatD are readily interconvertible, and that the inactive form may predominate in a variety of physiological conditions in

both lysosomal and extralysosomal compartments (Fig. 5). Expression of enzymatic activity requires interaction with substrates that can shift the equilibrium towards the active enzyme species, as evidenced by the stabilizing influence of pepstatin on $CatD_{lo}$. Thus, the specificity of CatD may be restricted by the ability of substrates to stabilize the low pH form of the enzyme as a prerequisite for substrate cleavage. This hypothesis may be relevant for other aspartic proteinases, such as renin, that are required to function at high pH and exhibit restricted substrate specificity.

The discovery that a novel and stable conformational species of CatD exists with a blocked active site at high pH raises new possibilities for the mechanism of pH-induced inactivation of aspartic proteinases. Our findings also suggest a hypothesis for the pH dependent modulation of substrate specificity in which high pH may restrict cleavage to those substrates with the highest affinity. The complex pattern of structural changes observed for the interconverting forms of $CatD_{lo}$ and $CatD_{hi}$ represents a novel example of a pH-dependent conformational switch in an aspartic proteinase.

Methods

Crystallization and structure determination. Native cathepsin D was purified from human liver[29]. Crystallization was carried out by the hanging drop method in a 6 µl droplet containing equal volumes of 10 mg ml^{-1} protein solution in water and reservoir solution. The best crystals were obtained at 22 °C from 2% v/v PEG 400, 2.0 M ammonium sulfate, 100 mM Sodium Hepes, pH 7.5. Crystals were transferred to artificial mother liquor containing additional 25% glycerol and flash-frozen in a stream of nitrogen at 190 K. Data were collected at 190 K on a MAR image plate using CuKα X-ray produced by a Rigaku RU200 rotating anode X-ray generator and processed with programs DENZO and SCALEPACK[30]. The data set contained 68,982 unique reflections out of a total of 197,513 measured reflections and was 73% complete to 2.5 Å resolution [$I > 2\sigma(I)$]. The overall R_{merge} was 10.8%. The crystals belong to space group of $P2_12_12$, with a = 139.9 Å, b = 136.5 Å, c = 139.6 Å. The asymmetric unit consists of four independent molecules and has a solvent content estimated at 70%. The structure was solved by molecular replacement using the coordinates of the hexagonal crystal structure of CatD[11] as the search model and refined to a 2.5 Å resolution with the program X-PLOR[31] without non-crystallographic symmetry restraints. The final model contains four protein monomers, four Hepes buffer molecules, and 470 water molecules with a crystallographic R-factor of 19.5% (R_{free} = 25.7%) for the 63,094 reflections >2σ in the resolution range 8.0–2.5 Å. The root-mean-square deviations for bond lengths and angles are 0.010 Å and 1.9°; the average B-values for protein and solvent molecules are 18.1 $Å^2$ and 17.4 $Å^2$ respectively. The final $2F_o - F_c$ map contoured at 1σ level showed well-defined, continuous electron density for almost all residues except for the first two N-terminal residues, Gly 1 and Pro 2, and the side chain of Gln 97 in all four monomers.

Enzyme activity, stability and ANS binding studies. The following buffers were used: glycine-HCl, pH 2.0–3.5; sodium acetate, pH 3.75–5.5; sodium-potassium phosphate, pH 6.0–8.0. The initial rate of fluorogenic substrate[32] cleavage in 0.2 M buffer was used for the determination of pH optima. For the pH stability study, 0.82 mg ml^{-1} CatD was incubated for 100 min, at 37 °C in 36 mM buffers, and the remaining activity was measured at pH 3.0 using hemoglobin as a substrate. For the ANS binding experiment, 0.2 mg ml^{-1} CatD was pre-incubated for 10 min in 0.2 M buffer with or without a two-fold molar excess of pepstatin. A 2.5 mM stock solution of ANS was then added to a final concentration of 25 µM and fluorescent emission spectra were recorded using a SLM-2 Luminescence spectrometer in the range 420–600 nm using an excitation wavelength of 360 nm. Baseline spectra were recorded for ANS in buffer with or without pepstatin and subtracted from the protein spectra. Fluorescence spectra at each pH were measured three times and the average spectrum was used for data analysis.

Coordinates. The coordinates of the $CatD_{hi}$ crystal structure have been deposited in the Brookhaven Protein Data Bank (accession number 1LYW).

Acknowledgments

We thank T.N. Bhat and A. Silva for crystallographic discussions and assistance in data collection, P. Majer for the fluorogenic substrate, D. Xie for the suggestion of the ANS experiments and B. Dunn for helpful discussions. Supported by NCI Contract No. NO1-CO-56000.

Angela Y. Lee[1,2], Sergei V. Gulnik[1,2] and John W. Erickson[1]

[1]Structural Biochemistry Program, SAIC Frederick, National Cancer Institute, P.O. Box B, Frederick, Maryland 21702-1201, USA. [2]These authors contributed equally to this work.

Correspondence should be addressed to J.W.E. *email: erickson@ncifcrf.gov*

Received 1 June, 1998; accepted 4 August, 1998.

1. Davies, D. R. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 189–215 (1990).
2. Tang, J. & Wong R. N. S. *J. Cell. Biochem.* **33**, 53–63 (1987).
3. Peters, P. J., Neefjes, J. J., Oorschot, V., Ploegh, H. L. & Geuze, H. J. *Nature* **349**, 669–676 (1991).
4. Bankowska, A., Gacko, M., Chyczewska, E. & Worowska, A. *Rocz. Akad. Med. Bialymst* . **42**, Suppl 1: 79–85 (1997).
5. Dreyer, R. N. *et al. Eur. J. Biochem*, **224**, 265–271 (1994).
6. Ladror, U. S., Snyder, S. W., Wang, G. T., Holzman, T. F. & Krafft, G. A. *J. Biol. Chem.* **269**, 18422–18428 (1994).
7. Garcia, M. *et al. Stem Cells* **14**, 642–50 (1996).
8. Saftig, P. *et al. EMBO J.* **14**, 3599–3608 (1995).
9. Majer, P., Collins, J. R., Gulnik, S. V. & Erickson, J. W. *Protein Sci.* **6**, 1458–1466 (1997).
10. Kick, E. K. *et al. Chem. & Biol.* **4**, 297–307 (1997).
11. Baldwin, E. T. *et al. Proc. Natl. Acad. Sci. USA* **90**, 6796–6800 (1993).
12. Metcalf, P. & Fusek, M. *EMBO J.* **12**, 1293–302 (1993).
13. Wlodawer, A. & Erickson, J. W. *Annu. Rev. Biochem.* **62**, 543–583 (1993).
14. Sail, A. *et al. Proteins. Struct. Funct. Genet.* **12**, 158–170 (1992).
15. Silva, A. M. *et al. Proc. Natl. Acad. Sci. USA*, **93**, 10034–10039 (1996).
16. Erickson, J. W., Baldwin, E. T., Bhat, T. N. & Gulnik, S. *In Aspartic Proteinases: Structure, Function, Biology and Biomedical Implications,* (ed. Takahashi, K.), 181 –192 (Plenum, New York, 1995).
17. James, M. N. G. & Sielecki, A. R. *Nature* **319**, 33–38 (1986).
18. Hartsuck, J. A., Koelsch, G. & Remington, S. J. *Proteins Struct. Funct. Genet.* **13**, 1–25 (1992).
19. Moore, S. A., Sielecki, A. R., Chernaia, M. M., Tarasova, N. I. & James, M. N. G. *J. Mol. Biol.* **247**, 466–485 (1995).
20. Khan, A. R., Cherney, M. M., Tarasova, N. I. & James, M. N. G. *Nature Struct. Biol.*, **4**, 1010–1015 (1997).
21. Fusek, M., Mares, M., Vagner, J., Vokurka, Z. & Baudys, M. *FEBS Lett.* **287**, 160–162 (1991).
22. Richards, A. D. *J. Biochem.* **24**, 297–301 (1992).
23. Dunn, B. M. *et al. Aspartic proteinases and their inhibitors*, (ed. Kostka, V.), 221-243 (Walter de Gruyter, Berline, New York, 1985).
24. Wittlin, S., Rosel, J. & Stover, D. R. *Eur. J. Biochem.* 252, 530–536 (1998).
25. Dunn, B. *Nature Struct. Biol.* 4, 967–972 (1997).
26. Stryer, L. *J. Mol. Biol.* **13,** 482–495 (1965).
27. Bohley, P. & Seglen, P. O. *Experientia*, 48, 151–157 (1992).
28. Tsuboi, M. *et al. J Appl. Physiol* .74, 1628–1634 (1993).
29. Gulnik, S., Baldwin, E. T., Tarasova, N. & Erickson, J. W. *J. Mol. Biol.* 227, 265–270 (1992).
30. Otwinowski, Z. *SCALEPACK*, (Yale Univ. Press, New Haven, CT, 1993).
31. Brünger, A. T. *X-PLOR, Version 3.1, A system for X-ray crystallography and NMR* (Yale Univeristy, New Haven, CT, 1992).
32. Gulnik, S. V. *et al. FEBS Letters* **413**, 379–384 (1997).
33. Rossmann, M. G. & Argos, P. *J. Biol. Chem.* **250**, 7525–7532 (1975).
34. Carson, M. *J. Appl. Crystallogr.* **24**, 958–961 (1991).
35. Nicholls, A., Bharadwaj, R. & Honig, B. *Biophys, J.* **64**, A166 (1993).