Recombinant Human Cathepsin H Lacking the Mini Chain Is an Endopeptidase†

Olga Vasiljeva,‡,§ Marko Dolinar,‡ Vito Turk,‡ and Boris Turk*,‡

Department of Biochemistry and Molecular Biology, Josef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia, and Department of Biochemistry and Molecular Biology, Siberian State Medical University, Tomsk, Russia

Received July 31, 2003; Revised Manuscript Received September 19, 2003

ABSTRACT: Human procathepsin H was expressed in the form of inclusion bodies in *Escherichia coli*. Following refolding and autocatalytic activation, a recombinant cathepsin H form lacking the mini chain was produced. Removal of the mini chain completely abolished aminopeptidase activity of the enzyme and largely increased its endopeptidase activity (∼40-fold). Similarly to cathepsin S, Bz-FVR-AMC (k_{cat}/K_m value of 1070 mM^{-1} s^{-1}) was found to be the preferred substrate of recombinant cathepsin H. However, substrate inhibition was observed at a higher substrate (Z-FR-AMC, Bz-FVR-AMC) concentration. Endopeptidase activity of recombinant cathepsin H was seen also with the protein substrate insulin β-chain with the major cleavage site between Glu13-Ala14. Recombinant human cathepsin H was inhibited by chicken cystatin, stefin A, and stefin B with the K_i values in the range of 0.05−0.1 nM, which is slightly tighter than the inhibition of purified cathepsin H by the same inhibitors. These results thus indicate that the cathepsin H mini chain is essential for the aminopeptidase activity of the enzyme but has only a minor effect on the inhibition by cystatins.

Cathepsin H (EC 3.4.22.16) is a ubiquitously expressed member of the papain family of cysteine lysosomal proteinases, which plays an important role in physiological intracellular protein degradation (*1*, *2*). The enzyme was also found to be connected with various pathologies, such as cancer and conditional mucolipidosis II (*3*).

Cathepsin H is unique among lysosomal cysteine proteases in that it is both an aminopeptidase and an endopeptidase (*1*, *4*, *5*), although the latter activity is much lower than the former activity (*5*, *6*). Sequencing data revealed that in addition to the heavy and light chains, which are typically found in a number of mammalian papain-like cysteine proteases, cathepsin H contained also an octapeptide originating from the propeptide, termed the mini chain (*7*, *8*). Further studies revealed that the mini chain is disulfide linked to Cys205 of the main body of the enzyme, where it has been suggested to play a role in the aminopeptidase activity of the enzyme (*9*). The crystal structure of the porcine enzyme (*10*) has confirmed the hypothesis and revealed that the mini chain fills the active site cleft of cathepsin H in the region equivalent to the S2 and S3 binding sites in the related endopeptidases. Thereby, the mini chain prevents access of substrates into these nonprimed subsites, whereas its C-terminal carboxyl group provides the negative charge required for the docking of the positively charged N-terminus of the substrate.

The K_i values in the nanomolar range were determined for inhibition of cathepsin H by cystatins, the endogenous inhibitors of lysosomal cysteine proteases (*11*−*14*). This is in contrast to endopeptidases, which are inhibited in the picomolar range (*15*). The mini chain, as the obvious additional structure feature, was therefore suggested to interfere with inhibitor binding (*10*, *12*).

We produced recombinant procathepsin H in the *Escherichia coli* expression system, which was autocatalytically activated at acidic pH. This resulted in production of cathepsin H lacking the mini chain, which enabled us to investigate the role of the mini chain in the functioning of cathepsin H by studying its interaction with various synthetic and protein substrates and protein inhibitors.

EXPERIMENTAL PROCEDURES

Materials. Restriction endonucleases and DNA-modifying enzymes were obtained from New England BioLabs (Beverly, MA) and Amersham Pharmacia Biotech (Uppsala, Sweden). Bz-Phe-Val-Arg-AMC (7-amido-4-methylcoumarin), Z-Phe-Arg-AMC, Z-Arg-Arg-AMC, and H-Arg-AMC were purchased from Bachem (Bubendorf, Switzerland), whereas E-64[1] [*trans*-epoxysuccinyl-L-leucyl-amido-(4-guanidino)butane] was from Peptide Research Foundation (Osaka, Japan). Iodoacetic acid, poly(ethylene glycol) 6000, and insulin β-chain were obtained from Serva (Heidelberg, Germany), and dimethyl sulfoxide was from Merck (Darmstadt, Germany). Sequencing reagents were from Applied Biosystems (Foster City, CA). Porcine spleen cathepsin H was isolated and purified according to ref *13*. Chicken cystatin and recombinant human stefin A and stefin B were prepared as described previously (*16*, *17*).

Production of Recombinant Human Procathepsin H. The expression vector for the recombinant human pcatH was

† This work was supported by the Ministry of School, Science and Sports of the Republic of Slovenia.

* Corresponding author. Tel.: +386 1 477 37 72. Fax: +386 1 257 35 94. E-mail: boris.turk@ijs.si.

‡ Josef Stefan Institute.

§ Siberian State Medical University.

[1] Abbreviations: AMC, 7-amido-4-methylcoumarin; Bz, benzoyl; E-64, L-*trans*-epoxysuccinyl-Leu-amido-(4-guanidino)butane; np cat H, native porcine cathepsin H; rh cat H, recombinant human cathepsin H; Z, benzyloxycarbonyl.

10.1021/bi035355k CCC: $25.00
Published on Web 10/30/2003

prepared essentially as described earlier (*18*). Briefly, mRNA was isolated from human uteral endometrium and converted to cDNA. The procathepsin H coding region was amplified by PCR. The primers used were 5′-AGGAAAGAATTCATATGGCCGAACTGTCCGTGAACTCCTTA-3′ and 5′-AATCTAGAAAGCTTGGATCCCGGGACGGCTCACACCAGAGGGA-3′. The 5′-amplification primer was designed to include restriction sites *Eco*RI and *Nde*I with the starting Met codon, followed by 24 bases complementary to the coding region for procathepsin H. The PCR product was inserted into pALTER, and the DNA sequence was checked using universal and internal sequencing primers. The expression vector pET3a was digested by *Bam*HI, and ends were filled with Klenow enzyme and then cleaved by *Nde*I. The procathepsin H coding fragment was excised from the pALTER vector using *Nde*I and *Sma*I and ligated into the expression vector. Expression plasmid, named pBMF2, was used for the transformation of *E. coli* BL21[DE3]pLysS. Recombinant human procathepsin H was then produced intracellularly and deposited as inclusion bodies, which were isolated, solubilized, and refolded essentially as described previously for procathepsins S and L (*19*, *20*).

Protein Concentration. Protein concentrations were determined by absorbance measurements at 280 nm (Perkin-Elmer λ-18 spectrometer, Wellesley, MA) using the molar absorption coefficients of 72685 and 35225 for procathepsin H and active cathepsin H, respectively, which were calculated by the method of Pace et al. (*21*). The active concentration of activated cathepsin H was determined by active site titration with stefin A, titrated with E-64-titrated papain (*22*).

Processing and Activation of Procathepsin H. Procathepsin H was processed and activated by acidification to pH 5.0 using 1.0 M Na-acetate, pH 4.0, and incubation at 37 °C with 2.5 mM dithiothreitol in the absence or presence of dextran sulfate at a final concentration of 25 μg/mL. This concentration was found to be optimal for the processing of procathepsin H (not shown) and is the same as previously used for cathepsins L, S, and B (*19*, *20*, *23*). Procathepsin H processing was analyzed by SDS−PAGE under reducing conditions. The 12 μL aliquots of the reaction mixture (see previously) without dextran sulfate were taken every 30 min, and processing was stopped by the addition of 4 μL of 4 × SDS−PAGE loading buffer, containing 4% (v/v) of β-mercaptoethanol and 1% SDS, followed by 5 min of boiling. Protein bands were visualized by Coomassie brilliant blue staining.

The kinetics of activation was followed essentially as described previously for procathepsin B (*23*). Briefly, procathepsin H (1−5 μM) was acidified to pH 5.0 in a final volume of 0.5 mL. The 5 μL aliquots were then taken from the reaction mixture at appropriate times and mixed with 2.5 mL of substrate solution (10 μM Bz-FVR-AMC in 0.1 M phosphate buffer, pH 6.5, containing 1 mM EDTA and 0.1% (w/v) poly(ethylene glycol) 6000 (PEG) to prevent adsorption to the cuvette walls). Fluorescence of the released AMC was then monitored continuously for 1 min in a Perkin-Elmer LS-50 spectrofluorimeter (Norwalk, CT) at excitation and emission wavelengths of 370 and 460 nm, respectively.

Substrate Kinetics. Recombinant human cathepsin H and purified porcine cathepsin H were tested against three substrates: H-R-AMC as a standard aminopeptidase substrate for native cathepsin H and two endopeptidase substrates, Z-FR-AMC and Bz-FVR-AMC, normally used for related cathepsins (*5*). The kinetics of substrate hydrolysis was monitored at 25 °C in 0.1 M phosphate buffer, pH 6.5, containing 1 mM EDTA and 0.1% (v/v) PEG as described previously (*19*, *24*). The concentration of dimethyl sulfoxide in the assay was adjusted to 4.5% (v/v).

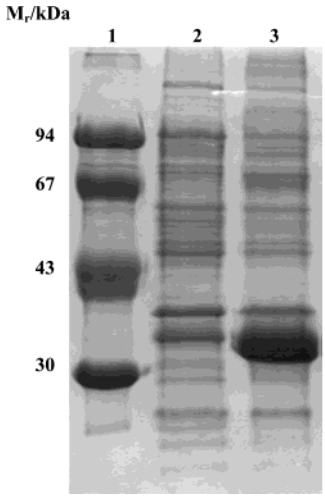

FIGURE 1: SDS-PAGE of recombinant procathepsin H biosynthesis. Lane 1, low molecular weight standards; lane 2, total cell extracts of bacteria BL21[DE3]pLysS containing pBMF2 before induction; and lane 3, total cell extracts of the same bacteria 3.5 h after addition of the inducer (IPTG). Protein bands were visualized by Coomassie brilliant blue staining.

Inhibition Kinetics. The kinetics of inhibition of recombinant cathepsin H by protein inhibitors chicken cystatin and stefins A and B were monitored for 15−30 min under pseudo-first-order conditions with at least a 10-fold molar excess of the inhibitors, essentially as described previously (*12*, *22*). The buffer used was the same as described previously, and 10 μM Bz-FVR-AMC was used as a substrate.

Insulin β-Chain Degradation. Insulin β-chain and recombinant cathepsin H were incubated at a molar ratio of 20:1 in acetate buffer, pH 5.0, containing 1 mM EDTA and 2.5 mM dithiothreitol at 37 °C. Aliquots of the reaction mixture were taken after 15 and 30 min, respectively, and the reaction was stopped immediately by the addition of 5 μL of 100 mM iodoacetate. The resulting peptides were then separated by HPLC (Milton Roy, Ivyland, PA) on a reverse phase Vydac C18 column (The Separations Group, CA), as described previously (*25*).

N-Terminal Protein Sequence Determination. The N-terminal sequences of peptides and proteins were determined using an Applied Biosystems (Foster City, CA) PROCISE 492A protein sequencing system.

RESULTS

Production of Recombinant Procathepsin H. Large quantities of recombinant procathepsin H were produced in *E. coli*. The recombinant protein was present predominantly in the insoluble fraction of the bacterial lysate and contributed to approximately 30% of total bacterial proteins as determined by the SDS−PAGE of cell extracts (Figure 1). After washing the inclusion bodies and the gel filtration step, which removed most of the impurities, ∼200 mg of procathepsin H with >90% purity as assessed by SDS−PAGE data (not

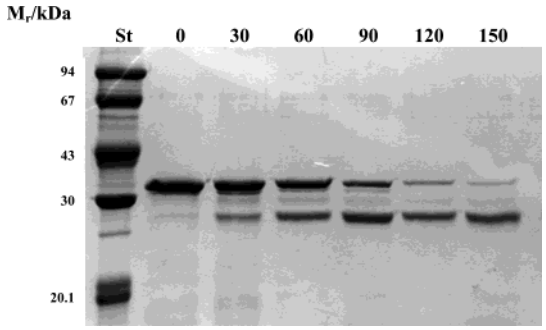

FIGURE 2: SDS-PAGE analysis of procathepsin H processing at pH 5.0 and 37 °C. Procathepsin H was incubated in the processing buffer as described under the Experimental Procedures in the absence of dextran sulfate. Samples were analyzed every 30 min. St; low molecular weight standards. The incubation times were as indicated: 0, 30, 60, 90, 120, and 150 min.

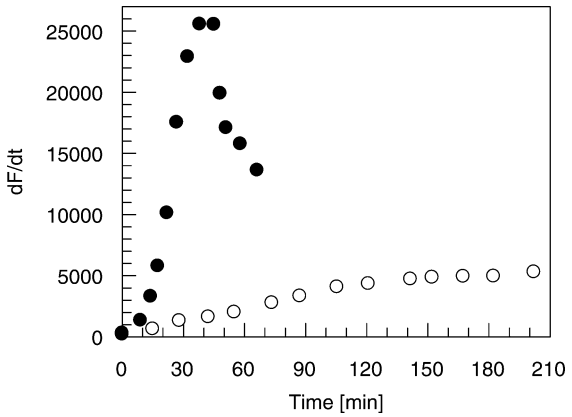

FIGURE 3: Autocatalytic processing of recombinant human procathepsin H at pH 5.0 and 37 °C. Processing mixtures were incubated in the absence (○) or presence (●) of 25 μg/mL dextran sulfate. Proteolytic activity was evaluated using 10 μM Bz-FVR-AMC as a substrate.

shown) was obtained from 4 L of bacterial culture. During refolding, which was a critical step, 95% of protein was lost resulting in a final yield of ~10 mg of properly folded procathepsin H.

Processing of Procathepsin H. When recombinant human procathepsin H was acidified to pH 5.0, a decrease in M_r corresponding to the loss of propeptide was observed, suggesting that procathepsin H has undergone autocatalytic processing. As shown in Figure 2, processing was completed in ~3 h. However, intensity of the band of procathepsin H prior to processing was substantially higher than the intensity of the band of cathepsin H after processing, suggesting that the protein has been at least partially degraded during processing. To check whether the processed recombinant human cathepsin H (rh cat H) was also activated, the activity of the enzyme was followed using Bz-FVR-AMC as a substrate since no activity was observed using the aminopeptidase substrate R-AMC. A sigmoidal increase of activity was observed with a maximum at ~180 min, in agreement with the electrophoresis data. In another experiment, activation was followed in the presence of 25 μg/mL of dextran sulfate, which has been shown to accelerate activation of cathepsins L and B (*23*, *26*, *27*). Activation was completed within 40−60 min, which was 3−4-fold faster. Furthermore, the final activity was much higher in the presence of dextran sulfate (Figure 3). However, prolonged activation resulted in the decrease of enzyme activity. After 120 min, the activity was <50% of the peak value and still declining (Figure 3), consistent with protein degradation (see previously). Therefore, for further studies, procathepsin H was activated for 40 min, concentrated at +4 °C, aliquoted, and stored at −70 °C.

N-Terminal Amino Acid Sequence Determination. To identify the N-termini of pro and active forms of rh cat H, samples were subjected to SDS−PAGE under nonreducing conditions and blotted to the PVDF membrane, followed by N-terminal amino acid sequencing. A single sequence starting with AELSV was obtained for procathepsin H, which is identical to the amino acid sequence deduced on the basis of the cDNA sequence (*28*). Similarly, a single sequence was obtained for the active form of rh cat H starting with GTGPYPPSVD. However, when the sequence was compared with the published sequence of human kidney cathepsin H (*8*) the sequence of rh cat H was four residues longer (GTGP). Moreover, N-terminal sequencing of human cathepsin H (*8*) revealed four N-terminal sequences, which were resolved into two forms of heavy chain, the major being two residues shorter than the minor, the light chain and the mini chain octapeptide, which suggested that the active rh cat H is a single chain enzyme lacking the mini chain. To confirm this result, the whole experiment was repeated with purified porcine cathepsin H. N-terminal sequencing of the blotted sample revealed four N-terminal sequences, starting with EPQXXSA corresponding to the mini chain, YPPSVDW corresponding to form 1 of the heavy chain, GPYPXSV corresponding to form 2 of the heavy chain, and GIPYWIV corresponding to the light chain of cathepsin H, consistent with the sequencing data of the purified human enzyme (*8*). In an additional experiment, SDS−PAGE and blotting were replaced by purification of pro and active forms of rh cat H by HPLC (reverse phase Vydac C4 column, The Separations Group, CA). N-terminal sequencing of both samples gave the same results as stated previously, indicating that the active form of rh cat H is indeed a single chain enzyme lacking the mini chain. However, the production of single chain recombinant cathepsins with N-terminal extensions is quite common and does not interfere with catalytic properties of the enzymes since none of the two features is positioned in the vicinity of the active site (reviewed in refs *1* and *29*).

Characterization of Cathepsin H Specificity. Initially, rh cat H was tested against the aminopeptidase substrate R-AMC. As expected, the rate of hydrolysis by rh cat H was found to be extremely low, when compared with native porcine cathepsin H (np cat H), suggesting that the recombinant enzyme was no longer an aminopeptidase. When the two enzymes were tested on two endopeptidase substrates, Bz-FVR-AMC and Z-FR-AMC, the pattern was reversed, and the recombinant enzyme was much more efficient than the native enzyme (~15−80-fold), suggesting that the recombinant enzyme has been converted to an endopeptidase (Figure 4).

In the next step, we wanted to determine the kinetic parameters for hydrolysis of the substrates by rh cat H. Surprisingly, the initial rates of substrate hydrolysis showed a substantial decrease at higher substrate concentrations for both Bz-FVR-AMC ($[S_o] \geq 30$ μM) and Z-FR-AMC ($[S_o] \geq 100$ μM). A similar effect has already been observed previously for cruzipain, a related enzyme from the parasite *Trypanosoma cruzi* (*24*), but not for other cathepsins. This

Table 1: Kinetic Constants for Cathepsin H Hydrolysis of Synthetic Substrates[a]

	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m[b] ($mM^{-1}\ s^{-1}$)	k_{cat}/K_m[c] ($mM^{-1}\ s^{-1}$)
		Recombinant Cathepsin H		
Bz-Phe-Val-Arg-AMC	38 ± 7.7	110 ± 13	2900 ± 930	1070 ± 36
Z-Phe-Arg-AMC	82 ± 26	4.3 ± 1.0	52.5 ± 28.8	37.5 ± 1.3
H-Arg-AMC				<5
		Native Cathepsin H		
Bz-Phe-Val-Arg-AMC (*48*)	25	1.6		64
Z-Phe-Arg-AMC (*48*)	nc	nc		nc
H-Arg-AMC (*54*)	100	8.0		80

[a] The experimental conditions are described in the Materials and Methods. [b] k_{cat}/K_m values were calculated from k_{cat} and K_m values obtained from the best fit of eq 1 (Figure 5). [c] k_{cat}/K_m values determined by linear regression analysis of the initial part of the same curves. All the values are given together with their standard errors. nc = not cleaved.

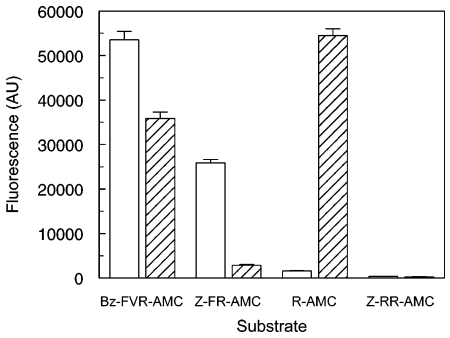

FIGURE 4: Substrate specificity of recombinant human and native porcine cathepsins H at pH 6.5 and 37 °C. Rh cat H (empty bars) and np cat H (hatched bars) were tested for their specificity against Bz-FVR-AMC, Z-FR-AMC, R-AMC, and Z-RR-AMC as substrates. Final concentration of all the substrates was 10 μM, whereas final concentrations of rh cat H and np cat H were 1 and 10 nM, respectively. Other experimental details were as described in the Experimental Procedures.

Scheme 1

$$E + S \underset{k_{-1}}{\overset{k_{+1}}{\longleftrightarrow}} ES \xrightarrow{k_{+2}} E + P$$

$$+$$

$$S$$

$$k_{+3} \updownarrow k_{-3}$$

$$ES_2$$

effect can be explained in terms of substrate inhibition and can be described by Scheme 1 (*30*).

According to Scheme 1, the rate of substrate hydrolysis can be best described by the following equation:

$$v = \frac{V_{max}}{1 + \frac{K_m}{[S]} + \frac{[S]}{K_{si}}} \quad (1)$$

where V_{max}, K_m, and [S] represent the maximal velocity, Michaelis−Menten constant, and total substrate concentration, respectively, and K_{si} stands for the equilibrium constant for substrate inhibition. As can be seen in Figure 5, eq 1 could be well-fitted to the experimental data, suggesting that the proposed model for substrate inhibition is correct. In the control experiments, the absorbance spectra of 100 μM product (AMC) and substrates (Z-FR-AMC and Bz-FVR-AMC) were recorded between 370 and 460 nm. Neither

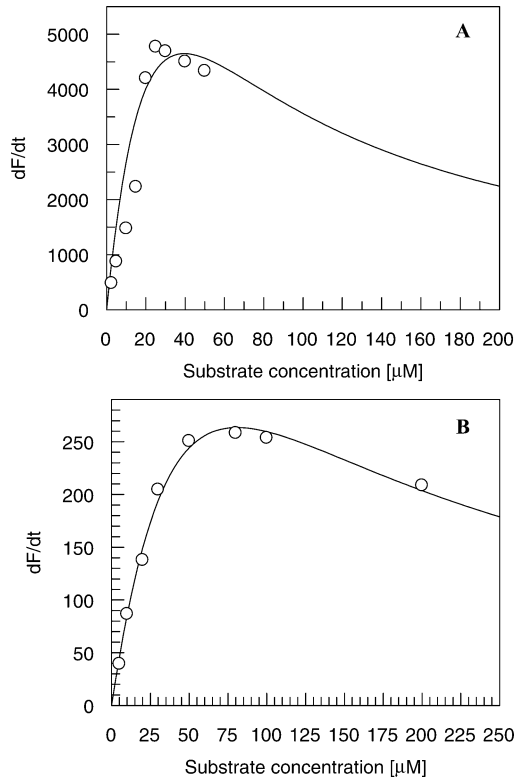

FIGURE 5: Substrate inhibition of recombinant cathepsin H at pH 6.5 and 25 °C. Substrate hydrolysis was followed fluorimetrically as described under the Experimental Procedures. Initial velocity rates are plotted as a function of substrate concentration. Solid lines represent the best fits of eq 1 to the experimental data. (A) Bz-FVR-AMC and (B) Z-FR-AMC.

AMC nor the substrates were found to absorb in this range, indicating that no correction for the inner-filter was needed and the effect was solely due to substrate inhibition. The K_m and k_{cat} values for the hydrolysis of both substrates (Z-FR-AMC and Bz-FVR-AMC) by rh cat H, obtained by this method, are given in Table 1. For comparison, the turnover number, k_{cat}/K_m, was also determined from the slope of the linear portion of the curve, which enables reliable determination of this parameter. Whereas the values for k_{cat}/K_m obtained by the two methods were in good agreement for Z-FR-AMC hydrolysis, they differ appreciably for Bz-FVR-AMC hydrolysis. The fit of eq 1 to the data for the latter substrate gave higher scattering. Hydrolysis of R-AMC by rh cat H was negligible, and only an upper estimate of the k_{cat}/K_m value is given (Table 1). A comparison of the k_{cat}/K_m values between the native and the recombinant cathepsin H has shown a ~40-fold increase for the Bz-FVR-AMC

Table 2: Inhibition Constants for the Interaction of Cathepsin H with Cystatins at pH 6.5 and 25 °C[a]

enzyme	inhibitor	$10^{-6} \times k_{ass}$ (M^{-1} s^{-1})	K_i (pM)	$10^4 \times k_{diss}$ (s^{-1})
recombinant human cathepsin H	human stefin A	4.0 ± 0.4	99 ± 14	3.9 ± 0.9
	human stefin B	3.5 ± 0.3	74 ± 5	2.6 ± 0.4
	chicken cystatin	1.0 ± 0.06	45 ± 3	0.45 ± 0.06
native human cathepsin H (*11*)	bovine stefin B	1.7	520	9.0
native bovine cathepsin H (*12*)	bovine stefin A	2.1	440	9.2

[a] Other experimental conditions were as described in the Materials and Methods. At least five different concentrations were used for each inhibitor. Values are given together with their standard errors.

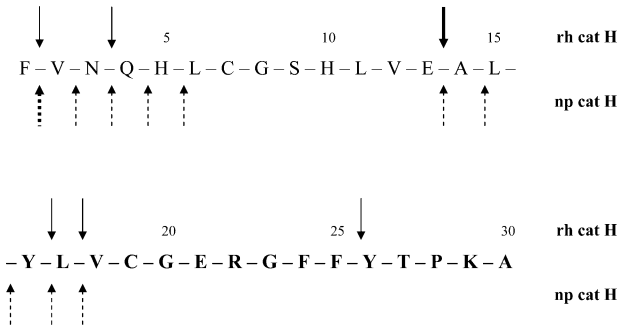

FIGURE 6: Cleavage of the oxidized insulin β chain by recombinant human (rh cat H) and native porcine (np cat H) cathepsins H. The cleavages determined after 30 min of incubation are shown. Cleavage sites are marked by full line arrows for rh cat H and by dashed arrows for np cat H. Major cleavage sites are marked by bold arrows.

hydrolysis and at least a 20-fold decrease for the hydrolysis of R-AMC by the latter, respectively. Comparison of other parameters was not possible due to substrate inhibition.

Hydrolysis of Insulin β-Chain by Recombinant and Native Cathepsin H. A further evaluation of endo- and exopeptidase activity of rh cat H was done using oxidized insulin β-chain as a substrate. Incubation of the insulin β-chain with both forms of cathepsin H resulted in substrate degradation. All the cleavage products were then separated by HPLC and subjected to N-terminal sequencing, which revealed the major cleavage site between Glu13-Ala14 and four minor cleavage sites for rh cat H (Figure 6). In contrast, the major cleavage site by np cat H, which was used as a control, was between Phe1-Val2. The other cleavages indicated sequential removal of single amino acid residues from the N-terminus, characteristic of an aminopeptidase. However, it should be noted that an internal cleavage site between Glu13-Ala14 has been observed for np cat H, indicating that the native enzyme is also an endopeptidase, in agreement with previous observations (*31*, *32*).

Inhibition Kinetics. The interaction between rh cat H and stefin A, stefin B, and chicken cystatin was monitored by continuous fluorimetric measurements. All progress curves were exponential with sloping endpoint and were analyzed by nonlinear regression analysis according to Morrison (*33*). Linear dependence of the observed pseudo-first-order rate constants on inhibitor concentrations were observed for all three inhibitors, in agreement with a simple, competitive reaction model (*33*). The k_{ass} values were calculated from the slopes of the plots. No correction for substrate competition was needed, as substrate concentration used in the experiment was well below the K_m value. The linearity of product formation in separate experiments without inhibitors verified that the enzyme was stable during the experiments.

The K_i values were calculated from the plots of $(v_z - v_s)/v_s$ versus $[I_o]$, with v_z representing the initial rate velocity and v_s the steady-state velocity (*22*). All three plots gave straight lines with a low degree of scattering. The values of k_{diss} were then calculated from the K_i and k_{ass} values ($k_{diss} = K_i k_{ass}$). As can be seen in Table 2, the K_i values for the interaction between all three inhibitors and rh cat H are very similar ($K_i = 45-100$ pM), as well as the corresponding k_{ass} values ($(1-5) \times 10^6$ M^{-1} s^{-1}). In the control experiment, performed in the absence of inhibitor, fluorescence of the released product increased linearly, verifying that the enzyme was stable during the experiment.

DISCUSSION

Crystal structures of the exopeptidases of papain-like lysosomal cysteine proteases revealed that the enzymes, in addition to the endopeptidase core, contain additional structural elements that restrict substrate access to the active site. Carboxypeptidases cathepsins B (*34*) and X (*35*) thus utilize loops, whereas aminopeptidases cathepsins C and H utilize propeptide parts: cathepsin C the exclusion domain (*36*), and cathepsin H is the mini chain (*10*). The same holds true also for the more distantly related cytoplasmic aminopeptidases bleomycin hydrolases, where the C-terminal extension has the same function (*37*−*39*).

Removal of these additional structural elements from the enzymes is therefore expected to result in a loss of exopeptidase activity and a major increase of endopeptidase activity. This has already been confirmed for cathepsin B (*40*), yeast bleomycin hydrolase (Gal6) (*38*), and PepC, a bacterial bleomycin hydrolase (*41*). Our results demonstrate that the cathepsin H form without the mini chain acts as an endopeptidase. Recombinant cathepsin H lacking the mini chain thus failed to degrade the standard cathepsin H substrate, R-AMC, and exhibited substantially increased activity against two other substrates, Z-FR-AMC and Bz-FVR-AMC, typically degraded by the related endopeptidases cathepsins S, L, V, K, and F. With a k_{cat}/K_m value of 1070 mM^{-1} s^{-1}, Bz-FVR-AMC was found to be a better substrate for rh cat H than Z-FR-AMC ($k_{cat}/K_m = 37.5$ mM^{-1} s^{-1}), which is the same pattern as observed for cathepsin S ($k_{cat}/K_m = 1605$ and 320 mM^{-1} s^{-1} for Bz-FVR-AMC and Z-FR-AMC, respectively; ref *5*). For cathepsins K, V, F, and L, the pattern is reversed, and Z-FR-AMC is hydrolyzed more efficiently (*5*, *42*−*44*). These results can be explained on a structural basis (Figure 7). In the crystal structure of np cat H, there is a small four residue loop (Lys155A−Asp155D) positioned just above the disulfide Cys200−Cys80P, by which the mini chain is attached to the body of the enzyme (numbering according to ref *10*). This loop is an insertion specific for cathepsin H and is not found in other papain-like cysteine proteases. It narrows the active site cleft and

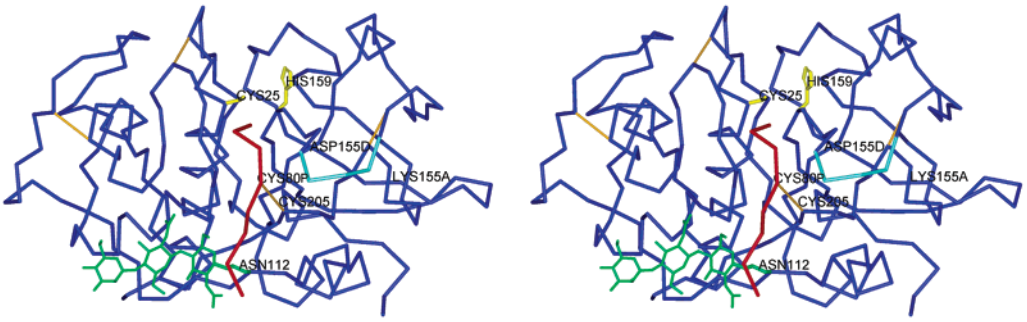

FIGURE 7: Stereo image of the crystal structure of porcine cathepsin H [10]. The mini chain (red) is positioned within the active site cleft (Cys 25 and Hys 159 are marked and are yellow) of the cathepsin H (blue). Carbohydrate moiety attached to Asn112 is shown in green, insertion loop specific for cathepsin H (Lys155A−Asp155D) in cyan, and the disulfide bonds (including the S−S bridge between the mini chain and the body of the enzyme) in orange. The figure was prepared with the ViewerLite 4.2 software (Accelrys, San Diego, CA).

stabilizes the positioning of the mini chain. In the complex of cathepsin H and stefin A (*45*), the loop is pushed away by the inhibitor N-terminus and disordered, thus exhibiting flexibility. In the absence of the mini chain, it is quite likely that the loop moves toward the S2 binding pocket, thereby limiting access of larger hydrophobic residues.

Endopeptidase activity of rh cat H was also confirmed by hydrolysis of the oxidized insulin β-chain, whereas np cat H acted primarily as an aminopeptidase. The cleavage pattern by recombinant enzyme was similar to those of cathepsins S (*46*) and L (*47*) with the major cleavage site between Glu13-Ala14, suggesting that cathepsin H without the mini chain behaves as a normal papain-like endopeptidase with substrate specificity similar to that of cathepsin S.

However, the same cleavage (Glu13−Ala14) has also been observed with purified porcine enzyme, although to a minor extent. This is in agreement with previous reports on the endopeptidase activity of native cathepsin H (*31*, *32*, *48*). From the crystal structure of the complex between cathepsin H and stefin A, it is evident that the mini chain of cathepsin H is flexible and can be moved from its normal position (*45*) to accommodate the N-terminus of the inhibitor. This flexibility of the mini chain might explain the binding of the larger substrates into the active site cleft of cathepsin H and the endopeptidase activity of the native enzyme. In this respect, cathepsin H seems to be similar to cathepsin B. The structural determinants of the exopeptidase activity, the occluding loop in cathepsin B and the mini chain in cathepsin H, are both flexible. In contrast, the mini loop of cathepsin X, the exclusion domain of cathepsin C, and the C-terminal insertion in bleomycin hydrolase and PepC appear to be more rigid, and cathepsins X and C, bleomycin hydrolase, and PepC act therefore only as exopeptidases.

This is reflected also in the activation mechanisms: whereas cathepsins X and C require other enzymes for their activation (*49*, *50*), cathepsins B and H can be autoactivated (refs *23* and *51*; Figures 2 and 3). We observed, however, that autocatalytic processing of cathepsin H resulted in a mutant lacking the mini chain, which is most likely a consequence of the *E. coli* expression system. From the crystal structure of np cat H, it is evident that there are two potential glycosylation sites in cathepsin H with at least Asn112 being glycosylated. Asn112 is located in the close vicinity of the disulfide bridge between the mini chain Cys80P and the main chain Cys205. Furthermore, the carbohydrate moiety was shown to interact with the mini chain (*10*). It is therefore possible that the carbohydrate moiety protects the disulfide Cys80P−Cys205 from reduction. Since the recombinant enzyme is nonglycosylated, it seems plausible that the Cys80P−Cys205 disulfide bond was not properly formed during refolding, thereby enabling the dissociation of the entire propeptide after the autocatalytic cleavage. This further suggests that the disulfide bridge has already formed in the zymogen. In vivo, however, there can also be other protease(s), such as cathepsin D, contributing to the processing (*52*).

As the mini chain was suggested to interfere also with protein inhibitor binding, inhibition of rh cat H by several cystatins has been investigated. All three inhibitors evaluated (chicken cystatin, human stefin A, and human stefin B) were shown to be tight-binding inhibitors of rh cat H with K_i values in the subnanomolar range (K_i = 0.045−0.1 nM). These values are somewhat lower than the values observed for the interaction between stefins and cystatins with purified cathepsin H, mainly due to the decreased k_{diss} values (*11*−*14*). This is in agreement with the proposed model for interaction between aminopeptidases cathepsins H and C and cystatins, where the octapeptide was suggested to interfere primarily in the second step, after the loop binding, locking the inhibitor to the active site by N-terminus binding (*53*). In the structure of the cathepsin H−stefin A complex, the N-terminus of the inhibitor was indeed shown to bend away from the active site due to steric clashes with the mini chain, forming an additional loop. In addition to the mini chain, the Lys155A−Asp155D loop moved from its position in the native enzyme, thereby narrowing the active site cleft (*45*). This can probably explain only a minor difference in affinity for cystatins between rh cat H and purified enzyme. To observe a larger effect, analogous to that observed on removal of the occluding loop in cathepsin B (*40*), the Lys155A−Asp155D loop would presumably have to be removed from cathepsin H as well.

In conclusion, a recombinant form of human cathepsin H lacking the mini chain was expressed in *E. coli*. The mutant enzyme was shown to be an endopeptidase, indicating that the mini chain is essential for the aminopeptidase activity of cathepsin H. However, the mini chain was shown to have a minor effect on protein inhibitor binding, indicating that an additional structural feature, presumably the insertion Lys155A−Asp155D specific for cathepsin H, narrows the active site cleft and prevents tighter binding of inhibitors.

ACKNOWLEDGMENT

We would like to thank Dušan Turk for fruitful dicussions, David Pim for critical reading of the manuscript, Ivica Štefe for the purified porcine cathepsin H, and Adrijana Leonardi for the N-terminal amino acid sequencing.

REFERENCES

1. Turk, B., Turk, D., and Turk, V. (2000) *Biochim. Biophys. Acta 1477*, 98−111.
2. Turk, V., Turk, B., and Turk, D. (2001) *EMBO J. 20*, 4629−4633.
3. Kopitz, J., Arnold, A., Meissner, T., and Cantz, M. (1993) *Biochem. J. 295*, 577−580.
4. Barrett, A. J., and Kirschke, H. (1981) *Methods Enzymol. 80*, 535−561.
5. Kirschke, H., Barrett, A. J., and Rawlings, N. D. (1995) Proteinases 1: lysosomal cysteine proteinases, *Protein Profile 2*, 1587−1643.
6. Mason, R. W. (1989) *Arch. Biochem. Biophys. 273*, 367−374.
7. Machleidt, W., Ritonja, A., Popovič, T., Kotnik, M., Brzin, J., Turk, V., Machleidt, I., and Müller-Esterl, W. (1986) in *Cysteine Proteinases and Their Inhibitors* (Turk, V., Ed.) pp 3−18, Walter de Gruyter, Berlin, New York.
8. Ritonja, A., Popovič, T., Kotnik, M., Machleidt, W., and Turk, V. (1988) *FEBS Lett. 228*, 341−345.
9. Baudyš, M., Meloun, B., Gan-Erdene, T., Fusek, M., Mareš, M., Kostka, V., Pohl, J., and Blake, C. C. (1991) *Biomed. Biochim. Acta 50*, 569−577.
10. Gunčar, G., Podobnik, M., Pungerčar, J., Štrukelj, B., Turk, V., and Turk, D. (1998) *Structure 6*, 51−61.
11. Pol, E., and Björk, J. (1999) *Biochemistry 38*, 10519−10526.
12. Turk, B., Ritonja, A., Björk, I., Stoka, V., Dolenc, I., and Turk, V. (1995) *FEBS Lett. 360*, 101−105.
13. Popovič, T., Brzin, J., Kos, J., Lenarčič, B., Machleidt, W., Ritonja, A., Hanada, K., and Turk, V. (1988) *Biol. Chem. Hoppe-Seyler 369*, 175−183.
14. Jerala, R., Kroon-Žitko, L., Popovič, T., and Turk, V. (1994) *Eur. J. Biochem. 224*, 797−802.
15. Turk, B., Turk, V., and Turk, D. (1997) *Biol. Chem. 378*, 141−150.
16. Kos, J., Dolinar, M., and Turk, V. (1992) *Agents Actions Suppl. 38*, 331−339.
17. Jerala, R., Trstenjak, M., Lenarčič, B., and Turk, V. (1988) *FEBS Lett. 239*, 41−44.
18. Dolinar, M., Mehle, A., Mozetič-Francky, B., Schweiger, A., and Turk, V. (2000) *Food Technol. Biotechnol. 38*, 5−9.
19. Kopitar, G., Dolinar, M., Štrukelj, B., Pungerčar, J., and Turk, V. (1996) *Eur. J. Biochem. 236*, 558−562.
20. Barlič-Maganja, D., Dolinar, M., and Turk, V. (1998) *Biol. Chem. 379*, 1449−1452.
21. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411−2423.
22. Turk, B., Križaj, I., Kralj, B., Dolenc, I., Popovič, T., Bieth, J. G., and Turk, V. (1993) *J. Biol. Chem. 268*, 7323−7329.
23. Rozman, J., Stojan, J., Kuhelj, R., Turk, V., and Turk, B. (1999) *FEBS Lett. 459*, 358−362.
24. Stoka, V., McKerrow, J. H., Cazzulo, J. J., and Turk, V. (1998) *FEBS Lett. 429*, 129−133.
25. Turk, B., Dolenc, I., Lenarčič, B., Križaj, I., Turk, V., Bieth, J. G., and Björk, I. (1999) *Eur. J. Biochem. 259*, 926−932.
26. Mason, R. W., and Massey, S. D. (1992) *Biochem. Biophys. Res. Commun. 189*, 1659−1666.
27. Tankersley, D. L., and Finlayson, J. S. (1984) *Biochemistry 23*, 273−279.
28. Fuchs, R., and Gassen, H. G. (1989) *Nucleic Acids Res. 17*, 9471.
29. Brömme, D., Nallaseth, F. S., and Turk, B. (2003) *Methods*, in press.
30. Dixon, M., and Webb, E. C. (1979) *Enzymes*, Academic Press, New York.
31. Koga, H., Mori, N., Yamada, H., Nishimura, Y., Tokuda, K., Kato, K., and Imoto, T. (1992) *Chem. Pharm. Bull. 40*, 965−970.
32. Lokshina, L. A., Lubkova, O. N., Gureeva, T. A., and Orekhovich, V. N. (1985) *Vopr. Med. Khim. 31*, 125−130.
33. Morrison, J. F. (1982) *Trends Biochem. Sci. 7*, 102−105.
34. Musil, D., Zučič, D., Turk, D., Engh, R. A., Mayr, I., Huber, R., Popovič, T., Turk, V., Towatari, T., and Katunuma, N. (1991) *EMBO J. 10*, 2321−2330.
35. Gunčar, G., Klemenčič, I., Turk, B., Turk, V., Karaoglanovic-Carmona, A., Juliano, L., and Turk, D. (2000) *Struct. Fold. Des. 8*, 305−313.
36. Turk, D., Janjič, V., Štern, I., Podobnik, M., Lamba, D., Dahl, S. W., Lauritzen, C., Pedersen, J., Turk, V., and Turk, B. (2001) *EMBO J. 20*, 6570−6582.
37. Joshua-Tor, L., Xu, H. E., Johnston, S. A., and Rees, D. C. (1995) *Science 269*, 945−950.
38. Zheng, W., Johnston, S. A., and Joshua-Tor, L. (1998) *Cell 93*, 103−109.
39. Mata, L., Erra-Pujada, M., Gripon, J. C., and Mistou, M. Y. (1997) *Biochem. J. 328*, 343−347.
40. Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T., and Mort, J. S. (1997) *J. Biol. Chem. 272*, 1197−1202.
41. Mata, L., Gripon, J. C., and Mistou, M. Y. (1999) *Protein Eng. 12*, 681−686.
42. Bossard, M. J., Tomaszek, T. A., Thompson, S. K., Amegadzie, B. Y., Hanning, C. R., Jones, C., Kurdyla, J. T., McNulty, D. E., Drake, F. H., Gowen, M., and Levy, M. A. (1996) *J. Biol. Chem. 271*, 12517−12524.
43. Brömme, D., Li, Z., Barnes, M., and Mehler, E. (1999) *Biochemistry 38*, 2377−2385.
44. Wang, B., Shi, G. P., Yao, P. M., Li, Z., Chapman, H. A., and Brömme, D. (1998) *J. Biol. Chem. 273*, 32000−32008.
45. Jenko, S., Dolenc, I., Gunčar, G., Doberšek, A., Podobnik, M., and Turk, D. (2003) *J. Mol. Biol. 326*, 875−885.
46. Brömme, D., Steinert, A., Friebe, S., Fittkau, S., Wiederanders, B., and Kirschke, H. (1989) *Biochem. J. 264*, 475−481.
47. Kargel, H. J., Dettmer, R., Etzold, G., Kirschke, H., Bohley, P., and Langner, J. (1980) *FEBS Lett. 114*, 257−260.
48. Xin, X.-Q., Gunesekera, B., and Mason, R. W. (1992) *Arch. Biochem. Biophys. 299*, 334−339.
49. Nägler, D. K., Tam, W., Storer, A. C., Krupa, J. C., Mort, J. S., and Menard R. (1999) *Biochemistry 38*, 4868−4874.
50. Dahl, S. W., Halkier, T., Lauritzen, C., Dolenc, I., Pedersen, J., Turk, V., and Turk, B. (2001) *Biochemistry 40*, 1671−1678.
51. Rowan, A. D., Mason, P., Mach, L., and Mort, J. S. (1992) *J. Biol. Chem. 267*, 15993−15999.
52. Nishimura, Y., and Kato, K. (1988) *Arch. Biochem. Biophys. 260*, 712−718.
53. Turk, B., Turk, D., and Salvesen, G. S. (2002) *Curr. Pharm. Des. 8*, 1623−1637.
54. Tchoupe, J. R., Moreau, T., Gauthier, F., and Bieth, J. G. (1991) *Biochem. Biophys. Acta 1076*, 149−151.

BI035355K