Amino acid substitutions in the N-terminal segment of cystatin C create selective protein inhibitors of lysosomal cysteine proteinases

Robert W. MASON^{*1}, Katia SOL-CHURCH^{*} and Magnus ABRAHAMSON[†]

^{*}Division of Developmental Biology, Nemours Research Programs, P.O. Box 269, Wilmington, DE 19899, U.S.A., and [†]Department of Clinical Chemistry, Institute of Laboratory Medicine, University of Lund, University Hospital, S-221 85 Lund, Sweden

We used site-directed mutagenesis to alter the specificity of human cystatin C, an inhibitor with a broad reactivity against cysteine proteinases. Nine cystatin C variants containing amino acid substitutions in the N-terminal (L9W, V10W, V10F and V10R) and/or the C-terminal (W106G) enzyme-binding regions were designed and produced in *Escherichia coli*. It was discovered that the inhibition profile of the cystatin could be altered by changing residues 9 and 10, which are proposed to bind in the S₃ and S₂ substrate-binding pockets respectively of the enzymes. All of the variants with substitutions in the N-terminal segment displayed decreased binding to cathepsins B and H, indicating that the S₃ and S₂ pockets of these enzymes cannot easily accommodate large aromatic residues. The introduction of a charged residue into S₂ (variant V10R) created a more specific inhibitor to distinguish cathepsin B from cathepsin H. Cathepsin L showed a preference for larger aromatic residues in S₂. In

contrast, cathepsin S preferred phenylalanine to valine in S₂, but bound less tightly to the V10W cystatin variant. The latter variant proved to be valuable for discriminating between cathepsin L and cathepsin S (K_i 2.4 and 190 pM respectively). The equilibrium dissociation constant of the complex between cathepsin L and variant L9W/W106G showed little difference in affinity from that of the cathepsin L complex with the singly substituted W106G variant. In contrast, the L9W/W106G variant displayed increased specificity for cathepsin S with a K_i of 10 pM. Our results clearly indicate differences in the specificity of interaction between the N-terminal region of cystatin C and cathepsins B, H, L and S, and that, although cystatin C has evolved to be a good inhibitor of all of the mammalian cysteine proteinases, more specific inhibitors of the individual enzymes can be engineered.

INTRODUCTION

Cystatin C is a natural inhibitor of the cysteine proteinase family represented by papain and the mammalian lysosomal cathepsins B, H, L and S. Although the mechanism by which cystatin C inhibits lysosomal cathepsins is not completely understood, three regions of its 120-residue polypeptide chain have been implicated in the formation of reversible enzyme–inhibitor complexes [1]. These regions, which are contiguous in the folded chain, form a wedge-shaped structure that can occlude the proteinases' substrate-binding sites, thus inhibiting them [2]. The first region in the N-terminal portion of cystatin C is characterized by the tripeptide sequence L⁹-V¹⁰-G¹¹ and is proposed to interact (in a substrate-like manner) with pockets S₃, S₂ and S₁ respectively of the cysteine proteinases. The evolutionarily conserved Q⁵⁵-I-V-A-G⁵⁹ and the C-terminal P¹⁰⁵-W¹⁰⁶ regions form two hairpin loop structures fitting closely in the S'₁ and S'₂ pockets of the enzyme, according to structural data for the avian cystatin C homologue, chicken cystatin [2].

The importance of the N-terminal region for the typical tight enzyme-binding property of cystatin C has previously been established [1,3,4]. We have shown that Leu-9 is the most discriminating residue for binding to cathepsins B, H, L and S, and thereby also a major determinant for the inhibitor's natural inhibition profile. Val-10 in this region does not confer specificity, unlike Leu-9, but is generally important because it makes a strong positive contribution to the binding of all these enzymes.

Our present goals were to determine whether cystatin C binds N-terminal residues in a manner analogous to that of synthetic

substrates and inhibitors, and to generate enzyme-specific protein inhibitors based on the cystatin C backbone. Short synthetic inhibitors and substrates have been used to show discriminating roles of residues present in P₂ and P₃ sites for the inhibition of cathepsins B, H, L and S ([5,6]; reviewed in [7]). On the basis of these results, we engineered recombinant cystatin C variants to study the effect of specific amino acid substitutions on their inhibitory properties. The recombinant proteins expressed in *Escherichia coli* were purified and used in kinetic experiments against the target proteinases, cathepsin B, H, L and S.

EXPERIMENTAL

Materials

Bovine spleens were purchased from Pelfreeze (Rogers, AR, U.S.A.). Synthetic fluorogenic substrates [Z-Phe-Arg-NHMec (where Z represents carbobenzyloxy and NHMec represents *N*-methylcoumaride), Z-Arg-Arg-NHMec, Arg-NHMec and benzoyl arginine *p*-nitroanilide (BAPNA)] were purchased from Calbiochem (La Jolla, CA, U.S.A.) and Sigma (St. Louis, MO, U.S.A.). Human cathepsin B and the cysteine proteinase inhibitor E-64 [*trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane] were purchased from Sigma. Papain, purchased from Sigma in a suspension at 25 mg/ml, was diluted with 0.01 % (v/v) Brij35 and stored in aliquots as a 0.25 mg/ml working solution. Restriction endonucleases and DNA-modifying enzymes were purchased from Life Technologies (Gaithersburg, MD, U.S.A.).

Abbreviations used: BAPNA, benzoyl arginine *p*-nitroanilide; E-64, *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane; NHMec, *N*-methylcoumaride; Z, carbobenzyloxy.

¹ To whom correspondence should be addressed.

Table 1 Oligonucleotides used for site-specific mutagenesis

The oligonucleotides shown (all minus strand, corresponding to the cystatin C cDNA sequence for residues 8–15) were synthesized to generate the desired N-terminal mutations by PCR as described in the text. Underlined sequences indicate variations from the wild-type cystatin C gene.

Mutation	Sequence
V10R	5'-GC GTC CAT GGG ACC ACC <u>ACG</u> CAG AC-3'
V10F	5'-GC GTC CAT GGG ACC ACC <u>AAA</u> CAG AC-3'
V10W	5'-GC GTC CAT GGG ACC ACC <u>CCA</u> CAG AC-3'
L9W	5'-GC GTC CAT GGG ACC ACC AAC <u>CCA</u> ACG CG-3'

Production of cystatin C variants

Mutagenic oligonucleotides were synthesized on an Applied Biosystems 392A synthesizer with phosphoramidites and other chemicals from Applied Biosystems (Foster City, CA, U.S.A.) (Table 1). Generation of the cystatin C variants with amino acid substitutions in the N-terminal region was performed with a PCR-based mutagenesis protocol described previously [4]. Plasmid pHD313 expressing wild-type cystatin C [8] was used to generate single mutants. Plasmid pCmut219 [4] containing the C-terminal W106G mutation was used to generate double mutants. Unique restriction sites for *ClaI* and *NcoI* were used to excise the coding region corresponding to the N-terminal region of the protein in the two parent plasmids, and these were exchanged for the mutated *ClaI*-*NcoI* fragment generated by PCR. Variants containing single substitutions are termed V10R, V10F, V10W, L9W and W106G and variants harbouring both N- and C-terminal substitutions are referred to as V10R/W106G, V10F/W106G, V10W/W106G and L9W/W106G.

The parent and variant plasmids were introduced into *E. coli* MC1061 [9] by CaCl_2 treatment [10]. Subclones of bacteria containing the expression plasmids were selected and plasmid DNA was isolated from the cultures by a standard alkaline lysis procedure [10]. All expression plasmid constructs were verified by DNA sequencing of the entire coding region of plasmids isolated from MC1061 subclones picked for expression, as described [4].

Production and purification of cystatin variants

Growth conditions for subclone cultures were exactly as described previously [11]. From a 1 litre culture of each subclone a 20 ml periplasmic extract containing the recombinant protein was obtained by cold osmotic shock. After overnight dialysis of the *E. coli* extract against 20 mM ethanolamine buffer, pH 9.0, the proteins were fractionated by FPLC on a Q-Sepharose column equilibrated in the same buffer. The unbound fractions were pooled and gel-fractionated on a Sephadex G-75 column equilibrated in 100 mM NH_4HCO_3 . The protein concentrations of the fractions containing inhibitory activities were determined by the bicinchoninic acid protein assay (Pierce, Rockford, IL, U.S.A.) and protein purity was confirmed by SDS/PAGE [16.5% (w/v) gel] and native electrophoresis at pH 8.6 in 1% (w/v) agarose gels [4].

Active-site titration of recombinant cystatins

Active-site titration of papain was done as described previously [12]. Stoichiometric titration of the recombinant proteinase inhibitors were performed in the presence of papain (previously active-site titrated with E-64) in a 1 ml reaction volume, with

BAPNA as chromogenic substrate. Briefly, 20 μl of papain (4 μM stock) was incubated at 37 °C for 30 min in 250 μl of activation buffer [400 mM sodium phosphate (pH 6.8)/4 mM EDTA/4 mM dithiothreitol] in the presence of a dilution series of inhibitors made up in 0.01% Brij35 to give a total incubation volume of 500 μl . An equal volume of 2.5 mM BAPNA made up in 0.01% Brij35 was added to each tube; the reaction proceeded for 30 min. BAPNA hydrolysis was stopped by adding 1 ml of 100 mM monochloroacetic acid/100 mM acetic acid (pH 4.3). Residual papain activity was recorded as absorbance at 410 nm. The A_{410} readings were plotted against the volumes of inhibitors. By extrapolating the initial linear part of the curve to the x-axis, the amount of recombinant inhibitor needed to inactivate papain was determined and the active site concentration calculated on the basis of a 1:1 stoichiometric reaction of enzyme and inhibitor.

Enzyme purification

Cathepsins S and H were prepared from bovine spleen as described previously [13]. Cathepsin L was purified from sheep liver [14]. The purity of each proteinase was determined by SDS/PAGE and staining with Coomassie Brilliant Blue.

Cystatin inhibition assays

The hydrolysis of methyl coumarylamide substrates was monitored continuously, as described previously [13], with a Perkin Elmer LS50 fluorimeter at an excitation wavelength of 360 nm and an emission wavelength of 460 nm. Briefly, cuvettes containing 750 μl of 400 mM activation buffer containing 4 mM dithiothreitol, 750 μl of 20 μM fluorogenic substrate and 1480 μl of 0.01% Brij35 were incubated for 10 min in thermostatically controlled holders. Cathepsins H and B (50 pM) were incubated at 37 °C in sodium acetate activation buffer, pH 6.0, using Arg-NHMec and Z-Phe-Arg-NHMec respectively as fluorogenic substrates. Cathepsins L (5 pM) and S (10 pM) were assayed with Z-Phe-Arg-NHMec. Cathepsin L was incubated at 30 °C in sodium acetate buffer, pH 5.5, whereas cathepsin S was incubated at 37 °C in sodium phosphate buffer, pH 7.0. When an initial steady-state (v_0) hydrolysis had been achieved, the recombinant inhibitor was added and reaction monitored until a new final steady-state rate (v_i) was reached. A range of inhibitor concentrations more than 10 times the enzyme concentrations that caused 20–80% enzyme inhibition were used in all cases except for the tightest-binding inhibitors. For the tightly binding cystatin C mutants that bound to cathepsin S, 3–10 times more inhibitor than enzyme was used to permit estimates of the low K_i values. For two of the cystatin variants that bound tightly to cathepsin L, more accurate determinations of K_i were made by using a high concentration of substrate (50 μM). This resulted in the measured apparent equilibrium constants, K_i (app), being much higher than the enzyme concentration, allowing a lower K_i to be calculated. K_i (app) values were calculated as described by Henderson [15] as the slope of the plot of $[I](1 - v_i/v_0)$ against v_0/v_i , in which $[I]$ is the inhibitor concentration in the assay, v_i is the inhibited activity and v_0 is the initial activity. K_i was calculated from these K_i (app) values by using the equation K_i (app) = $K_i(1 + [S]/K_m)$. K_m values for cathepsins B, H, L and S with corresponding substrates were 150, 40, 1.8 and 15 μM respectively [13,14,16]. These results were used to correct previously published K_i values that were derived from slightly different K_m values [4].

Reaction of cathepsin L with Z-Arg-Arg-NHMec

To determine whether cathepsin L could hydrolyse this compound, the enzyme was incubated with 20 μM Z-Arg-Arg-

NHMec in 100 mM sodium acetate buffer containing 1 mM EDTA and 1 mM dithiothreitol for up to 4 h. At 30 min intervals a portion of the incubation medium was applied to a Nova Pak C₁₈ column fitted to a Waters HPLC system. Absorbance was monitored at 325 nm and bound compounds were eluted with a linear gradient from 0.1 % trifluoroacetic acid in water to 100 % acetonitrile. The column was calibrated with Arg-NHMec, NHMec and Z-Arg-Arg-NHMec.

To determine the nature of the inhibition of cathepsin L activity by Z-Arg-Arg-NHMec, a series of continuous rate assays were set up as described above. When an initial steady-state hydrolysis of Z-Phe-Arg-NHMec (v_0) had been reached, Z-Arg-Arg-NHMec was added and the reaction monitored until a new final steady-state rate (v_i) was reached. The substrate (Z-Phe-Arg-NHMec) concentration in the assay ranged from 2 to 10 μ M and the Z-Arg-Arg-NHMec concentration ranged from 0 to 100 μ M. The inhibition profile and the K_i of the Z-Arg-Arg-NHMec-enzyme complex were determined by plotting $1/v$ against $[I]$ as described previously [17].

RESULTS AND DISCUSSION

Production and characterization of cystatin C variants

We have previously reported the construction of cystatin C variants by site-specific mutagenesis, and the expression of the recombinant proteins in bacteria [4]. The new series of single and double mutants analysed in this study were obtained by a similar procedure with similar yields of stable, functional inhibitors (see the Experimental section). The series comprised four novel variants with single amino acid substitutions in the N-terminal binding region: V10R-, V10F-, V10W- and L9W-cystatin C. The first mutant was chosen because of the known ability of cathepsin B, but not cathepsins H, L and S, to hydrolyse synthetic substrates with arginine residues in P₂. The other mutants were chosen because of the known preference of cathepsins S and L for synthetic substrates with hydrophobic residues in P₂ and P₃ (reviewed in [7]). A fifth variant in the series, W106G-cystatin C, was selected because the mutated tryptophan residue had been shown to contribute to the affinity between cystatin C and target enzymes independently of binding contributions from the N-terminal binding region [4]. Four variants carrying this substitution together with the N-terminal substitutions of the single mutants completed the series. The latter variants, V10R/W106G, V10F/W106G, V10W/W106G and L9W/W106G, should therefore be useful in studying the effects of the N-terminal substitutions for interactions with enzymes that bind too tightly to wild-type cystatin C to allow reliable equilibrium measurements in the continuous rate assays. All recombinant cystatin C variants were purified to homogeneity as judged by Coomassie staining after SDS/PAGE, and migrated as bands corresponding to a molecular mass of 14–15 kDa (Figure 1A). On non-denaturing agarose gels they also migrated as single entities, with V10R/W106G migrating more cathodally owing to the substitution of the positively charged amino acid (Figure 1B). Furthermore they all eluted from a calibrated gel-filtration column in accordance with the volume of monomeric cystatin C (results not shown). In addition to DNA sequencing of the entire coding regions of the expression vectors, all protein variants were subjected to N-terminal sequencing: twelve steps of Edman degradation for each verified the correct N-terminal sequence starting at residue Ser-1 and confirmed the intended substitutions. Methanesulphonic acid hydrolysis of the W106G-, (V10R/W106G)- and (V10F/W106G)-cystatin C mutants before quantitative amino acid analysis [4] confirmed the lack of tryptophan in these variants. Titration of each of the inhibitors

Figure 1 Electrophoresis of cystatin variants

The cystatin C variants with double substitutions were subjected to denaturing or native electrophoresis and then Coomassie stained. **(A)** Tricine/SDS/PAGE of samples after reduction with 2-mercaptoethanol with a separation gel containing 16.5% (w/v) polyacrylamide. The size markers shown in the flanking lanes are phosphorylase *b* (94 kDa), BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), soybean trypsin inhibitor (20.1 kDa) and α -lactalbumin (14.4 kDa). **(B)** Agarose gel (1%, w/v) electrophoresis at pH 8.6. A human blood plasma sample (M) was subjected to electrophoresis as a reference. The direction of the electric field is indicated. The point of sample application is marked by an arrowhead. Lanes 1 and 6, recombinant wild-type cystatin C; lane 2, (V10R/W106G)-cystatin C; lane 3, (V10F/W106G)-cystatin C; lane 4, (V10W/W106G)-cystatin C; lane 5, (L9W/W106G)-cystatin C.

Table 2 Equilibrium constants for dissociation, K_i , of complexes between mammalian cathepsins and recombinant variants of cystatin C

K_i determinations were made as described by Henderson [15]. At least five different concentrations of inhibitor were used for each determination. The S.D. for each point was typically less than 20%. Asterisks indicate results calculated with the use of 50 μ M substrate. Abbreviation: n.d., not determined.

Recombinant cystatin C	K_i (nM)			
	Cathepsin B	Cathepsin H	Cathepsin L	Cathepsin S
Wild-type	0.19	0.31	< 0.010	< 0.010
V10F	2.0	0.89	< 0.010	< 0.010
V10W	3.5	3.8	< 0.010	< 0.010
V10R	1.0	9.8	< 0.010	< 0.010
L9W	0.6	0.94	< 0.010	< 0.010
W106G	350	170	0.021	0.11
V10F/W106G	> 500	> 500	0.0016*	0.011
V10W/W106G	> 500	> 500	0.0024*	0.19
V10R/W106G	n.d.	n.d.	0.17	1.5
L9W/W106G	> 500	400	0.026	0.010

with papain confirmed that the purified proteins functioned as inhibitors.

Table 2 shows the equilibrium constants for dissociation (K_i) of the complexes between recombinant cystatin C and cathepsins

Figure 2 Relative affinities of cystatin C variants for cathepsins B, H, L and S

The K_i data derived in the present study (Table 2) and a previous one [4] were combined; affinities relative to the L9G (top panel) or V10G (bottom panel) mutant for each enzyme are shown. The single-letter code is used to signify amino acids in positions 9 and 10 for each enzyme. For cathepsins L and S, relative affinities were derived from cystatin C variants with an additional mutation of W106G. Relative affinities are presented on a logarithmic scale.

B, H, L and S. Figure 2 shows a comparison of affinities with the previously reported L9G and V10G mutants [4], respectively. As observed previously, recombinant wild-type cystatin C displayed a higher affinity for cathepsins L and S ($K_i < 10$ pM) than it did for cathepsins B and H, as seen by the higher K_i values for the latter enzymes.

W106G mutation

Substitution of the critical tryptophan residue of the C-terminal binding region of cystatin C had a destabilizing effect on the complex formations for all four cathepsins (Table 2). The affinities of binding to cathepsins B and H were decreased by approximately three orders of magnitude. Additional N-terminal substitutions nearly abolished the inhibitory activities of the variants for these two enzymes ($K_i > 500$ nM). The exact effect of the W106G mutation on binding to cathepsins L and S could not be determined because the binding of the wild-type inhibitor was too tight to measure under our experimental conditions. However, the weakest-binding double mutant (V10R/W106G) had binding affinities for cathepsins L and S that were respectively no more than 1/17 and 1/150 those of the corresponding single mutant (V10R), suggesting that the C-terminal binding region makes a major contribution to binding cystatin C to these enzymes. The binding to cathepsins L and S was weakened sufficiently by the W106G mutation to permit the determination of reliable K_i values for this and subsequent additional N-terminal mutations.

Position 9 mutations

We found previously that the leucine residue in this position provided the most discrimination between cathepsins B, H, L and S [4], and this can also be seen clearly in Figure 2 (top panel). Tryptophan in the 'P₃' site of cystatin C was even more poorly accommodated by cathepsin H than was leucine. However, a similar decrease in the affinity for cathepsin B was seen with the L9W mutant and thus this substitution did not enhance discrimination between these two enzymes. The preference of cathepsin B for the hydrophobic residue leucine in 'P₃' in contrast to the preference of cathepsin H for no side chain in 'P₃' results in the L9G mutant being a more tightly binding inhibitor of the latter enzyme by more than three orders of magnitude [4]. For cathepsin L, the L9W mutation was essentially neutral, whereas for cathepsin S it resulted in one order of magnitude tighter binding. In contrast with the W106G mutant, which binds cathepsin L five times more tightly than cathepsin S, the L9W/W106G mutant binds cathepsin S three times more tightly than cathepsin L.

Position 10 mutations

The V10F mutation produced an inhibitor whose binding to cathepsin B was one order of magnitude weaker than that of wild-type cystatin C and whose binding to cathepsin H was one-third as strong (Table 2 and Figure 2, bottom panel). When the bulk of the aromatic residue was increased by the V10W substitution, further decreases in affinity of binding to these two enzymes were seen. These results show that, although both cathepsins B and H have affinity for hydrophobic residues in P₂, they preferentially bind aliphatic side chains over aromatic ones. The V10R mutant was a better inhibitor of cathepsin B than any of the other mutants, although it was still a poorer inhibitor than the wild-type cystatin. By contrast V10R was as poor an inhibitor of cathepsin H as the V10G inhibitor, and thereby preferentially inhibited cathepsin B. The similar magnitude of inhibition of cathepsin B by the V10R and V10F mutants is consistent with the known ability of cathepsin B to bind and hydrolyse substrates that contain either of these residues in their S₂ subsite [16]. Cathepsin H has no known affinity for arginine in this subsite.

The V10F/W106G mutant exhibited 10-fold tighter binding to cathepsin L and cathepsin S than the corresponding W106G mutant. The V10W/W106G bound cathepsin L with an affinity similar to that of the V10F/W106G mutant, but bound much less tightly to cathepsin S. This resulted in the production of an inhibitor that was almost two orders of magnitude tighter in binding to cathepsin L than cathepsin S, enabling discrimination between these two enzymes. A 25 pM concentration of the V10F/W106G mutant would inhibit 90% of the activity of cathepsin L but have little effect on the activities of cathepsins B, H or S. Although both cathepsins L and S have a strong affinity for inhibitors with phenylalanine in 'P₂', our results suggest that the S₂ specificity pocket is more restricted in cathepsin S than in cathepsin L.

Previous results with low-molecular-mass synthetic substrates and inhibitors have provided conflicting conclusions on the specificity of cathepsin S for residues in P₂ [5,6]. It was proposed that cathepsin S preferred less bulky residues in P₂ when a series of substrates with a phenylalanine, leucine or valine residue at this site were studied [5]. By contrast, when a series of peptidyl vinyl sulphone inhibitors with a range of hydrophobic residues in P₂ was examined it was concluded that, unlike cathepsins L and K (a related cysteine proteinase), cathepsin S prefers bulkier groups in its S₂ subsite [6]. This conclusion was supported by the observation that residues in the S₂ subsite of cathepsin S are

smaller than those in cathepsins L and K, allowing more room for bulky residues. However, iodotyrosine in P₂ was shown to be a better inhibitor of cathepsin L than unmodified tyrosine in a peptidyl diazomethane inhibitor, supporting the conclusion that cathepsin L prefers bulky hydrophobic groups in S₂ [18]. Our results with the cystatin mutant that has a much bulkier tryptophan residue in P₂ also suggest that cathepsin L prefers bulky residues in S₂ and that cathepsin S tolerates them less well. Clearly size is not the only characteristic that determines the ability of a particular residue to bind to the subsite of a proteinase. The apparent discrepancies in conclusions reached with different molecular approaches might in part be due to the different modes of binding of the reagents and the contribution of residues in adjacent sites. Such interactions seem to be important between residues in P₁ and P'₁. Studies have shown that, whereas *N*-methylcoumarin reagents with arginine in P₁ are good substrates for cysteine proteinases, peptidyl diazomethanes with arginine in P₁ are poor inhibitors [19]. Furthermore *N*-methylcoumarin reagents with alanine in P₁ are poor substrates for cathepsin L [20], but peptidyl diazomethanes containing this residue in P₁ bind tightly [18]. These substrates and inhibitors all contained phenylalanine in P₂ but the substrates contained a bulky hydrophobic group in P'₁ (methylcoumarin), whereas the inhibitors contained a small diazomethane group in this position. Similar interactions might explain some of the differences in P₂ specificities seen when substrates that contain arginine in P₁ are compared with inhibitors containing homophenylalanine in P₁ [5,6]. In cystatins, glycine is in the P₁ position and our previous results indicated that residues in P₃ and P₂ interact independently with cysteine proteinases [4]. Unlike the synthetic compounds, cystatins are orientated by the two hairpin loop structures that fit in the S'₁ and S'₂ pockets of the enzymes and not the P₁ residue.

The substitution of an arginine residue in 'P₂' of cystatin C (V10R/W106G) resulted in an inhibitor that bound as weakly to cathepsin S as the corresponding glycine mutant (V10G/W106G), analogous to the result seen for cathepsin H. Like cathepsin H, cathepsin S has no known ability to bind or hydrolyse substrates with arginine in P₂. In contrast, although the V10R/W106G mutant was a poorer inhibitor of cathepsin L than the W106G mutant, it was still two orders of magnitude more tightly binding than the V10G/W106G mutant. This indicates that, in contrast with cathepsins H and S and like cathepsin B, cathepsin L can bind positively charged residues in S₂.

Binding of Z-Arg-Arg-NHMec by cathepsin L

The increased affinity of cathepsin L for the arginine mutant over the corresponding glycine mutant was surprising, considering the lack of evidence for cathepsin L's binding or hydrolysing substrates with arginine in P₂. We reinvestigated the ability of cathepsin L to hydrolyse Z-Arg-Arg-NHMec but were unable to detect any significant hydrolysis with this reagent, even at 100 μ M. HPLC analysis showed that Z-Arg-Arg-NHMec was not processed at all by the enzyme (results not shown). The enzyme also failed to hydrolyse Z-Arg-NHMec. For longer peptide substrates it has been shown that cathepsin L can cleave between two basic residues, although all of these contained phenylalanine in P₂ [17]. However, Z-Arg-Arg-NHMec did inhibit hydrolysis of Z-Phe-Arg-NHMec by cathepsin L. In a Dixon [21] plot, inhibition data generated at three different substrate concentrations gave three lines that intercept in the first quadrant at the same point (results not shown). These results showed that the inhibition observed is competitive, with a *K_i* of 10 μ M. Thus although cathepsin L cannot hydrolyse substrates with arginine in P₂ it is able to bind them. Presumably the

binding of Z-Arg-Arg-NHMec to cathepsin L does not result in the correct positioning for hydrolysis of the amide bond to the *N*-methylcoumarin by the enzyme.

The binding of Z-Arg-Arg-NHMec is only one-sixth as strong as that of Z-Phe-Arg-NHMec; the *K_m* of sheep cathepsin L for the Z-Phe-Arg-NHMec substrate is 1.8 μ M [14]. In contrast, the V10R/W106G mutant binds two orders of magnitude more weakly to cathepsin L than the V10F/W106G mutant. As discussed above, these differences might be due to the distinct influences of the residues occupying S₁ and S'₁ in cystatin C and the synthetic compound. These results clearly emphasize that although we have shown that the N-terminal portion of cystatin C binds to cathepsins B, H, L and S in a fashion similar to synthetic substrates, the binding modes are not identical.

In summary, a series of mutants of cystatin C have been developed that have discriminating affinities for the mammalian lysosomal cysteine proteinases, cathepsins B, H, L and S. Our results clearly show that the N-terminal segment of cystatin C interacts with the non-primed substrate pockets of cathepsins B, H, L and S in a manner similar to small-molecular-mass substrates and inhibitors. All of the mutants bind more tightly to cathepsins L and S than cathepsins B and H, presumably owing to the tight binding of the former enzymes to the QVIAG loop of cystatin C. The V10R mutation provides an inhibitor with specificity for cathepsin B over cathepsin H, and the V10W/W106G mutations provide an inhibitor with specificity for cathepsin L over cathepsin S. These results, coupled with previous results that showed that the L9G inhibitor has specificity for cathepsin H over cathepsin B and that the R8G/L9G/V10G/W106G inhibitor has specificity for cathepsin S over cathepsin L, demonstrate that enzyme-specific inhibitors can be engineered based on cystatin C. Although cystatin C is one of the best protein inhibitors of mammalian cysteine proteinases, our results demonstrate that it has not evolved to be an optimal inhibitor for any of the major mammalian cysteine proteinases except perhaps for cathepsin B.

We thank Darlise O'Connor, Katherine Pell, Brian Sander and Anne-Cathrine Löfström for technical assistance, and Charles Kettner for reminding us that high substrate concentrations can be used to enable low *K_i* values to be determined. This work was supported by the Nemours Foundation and grants from the National Science Foundation (grant no. MCB-9496260), the American Lung Association of Virginia, the Medical Faculty of the University of Lund, Magn. Bergvall's Foundation and the Swedish Medical Research Council (grant no. 09915).

REFERENCES

- 1 Abrahamson, M., Ritonja, A., Brown, M. A., Grubb, A., Machleidt, W. and Barrett, A. J. (1987) *J. Biol. Chem.* **262**, 9688–9694
- 2 Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J. and Turk, V. (1988) *EMBO J.* **7**, 2593–2599
- 3 Bjork, I., Pol, E., Rayb-Segall, E., Abrahamson, M., Rowan, A. D. and Mort, J. S. (1994) *Biochem. J.* **299**, 219–225
- 4 Hall, A., Hakansson, K., Mason, R. W., Grubb, A. and Abrahamson, M. (1995) *J. Biol. Chem.* **270**, 5115–5121
- 5 Bromme, D., Steinert, A., Friebe, S., Fittkau, S., Wiederanders, B. and Kirschke, H. (1989) *Biochem. J.* **264**, 475–481
- 6 Bromme, D., Klaus, J. L., Okamoto, K., Rasnick, D. and Palmer, J. T. (1996) *Biochem. J.* **315**, 85–89
- 7 Mason, R. W. and Wilcox, D. (1993) in *Advances in Cell and Molecular Biology of Membranes*, vol. 1 (Storrie, B. and Murphy, R., eds.), pp. 81–116, JAI Press, New York
- 8 Abrahamson, M., Dalboge, H., Olafsson, I., Carlsen, S. and Grubb, A. (1988) *FEBS Lett.* **236**, 14–18
- 9 Casadaban, M. J. and Cohen, S. H. (1980) *J. Mol. Biol.* **138**, 179–207
- 10 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

- 11 Hall, A., Dalboge, H., Grubb, A. and Abrahamson, M. (1993) *Biochem. J.* **291**, 123–129
- 12 Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirschke, H., Knight, G. C., Tamai, M. and Hanada, K. (1982) *Biochem. J.* **201**, 189–198
- 13 Xin, X. Q., Gunesequera, B. and Mason, R. W. (1992) *Arch. Biochem. Biophys.* **299**, 334–339
- 14 Mason, R. W. (1986) *Biochem. J.* **240**, 285–288
- 15 Henderson, P. J. F. (1972) *Biochem. J.* **127**, 321–333
- 16 Barrett, A. J. and Kirschke, H. (1981) *Methods Enzymol.* **80**, 535–561
- 17 Menard, R., Carmona, E., Plouffe, C., Bromme, D., Konishi, Y., Lefebvre, J. and Storer, A. C. (1993) *FEBS Lett.* **328**, 107–110
- 18 Crawford, C., Mason, R. W., Wikstrom, P. and Shaw, E. (1988) *Biochem. J.* **253**, 751–758
- 19 Zumbrunn, A., Stone, S. and Shaw, E. (1988) *Biochem. J.* **250**, 621–623
- 20 Mason, R. W., Green, G. D. J. and Barrett, A. J. (1985) *Biochem. J.* **226**, 233–241
- 21 Dixon, M. (1953) *Biochem. J.* **55**, 170–171

Received 13 August 1997/29 October 1997; accepted 21 November 1997