Grafting of Features of Cystatins C or B into the N-Terminal Region or Second Binding Loop of Cystatin A (Stefin A) Substantially Enhances Inhibition of Cysteine Proteinases†

Alona Pavlova and Ingemar Björk*

Department of Molecular Biosciences, Section of Veterinary Medical Biochemistry, Swedish University of Agricultural Sciences, Uppsala Biomedical Center, Box 575, SE−751 23 Uppsala, Sweden

Received May 12, 2003

ABSTRACT: Replacement of the three N-terminal residues preceding the conserved Gly of cystatin A by the corresponding 10-residue long segment of cystatin C increased the affinity of the inhibitor for the major lysosomal cysteine proteinase, cathepsin B, by ∼15-fold. This tighter binding was predominantly due to a higher overall association rate constant. Characterization of the interaction with an inactive Cys29 to Ala variant of cathepsin B indicated that the higher rate constant was a result of an increased ability of the N-terminal region of the chimeric inhibitor to promote displacement of the cathepsin B occluding loop in the second binding step. The low dissociation rate constant for the binding of cystatin A to cathepsin B was retained by the chimeric inhibitor, which therefore had a higher affinity for this enzyme than any natural cystatin identified so far. In contrast, the N-terminal substitution negligibly affected the ability of cystatin A to inhibit papain. However, substitutions of Gly75 in the second binding loop of cystatin A by Trp or His, making the loop similar to those of cystatins C or B, respectively, increased the affinity for papain by ∼10-fold. This enhanced affinity was due to both a higher association rate constant and a lower dissociation rate constant. Modeling of complexes between the two variants and papain indicated the possibility of favorable interactions being established between the substituting residues and the enzyme. The second-loop substitutions negligibly affected or moderately reduced the affinity for cathepsin B. Together, these results show that the inhibitory ability of cystatins can be substantially improved by protein engineering.

Cystatins are protein inhibitors of papain-like cysteine proteinases (proteinase family C1) (reviewed in refs *1*−*5*). Mammalian cystatins efficiently inhibit endogenous lysosomal cysteine proteinases, e.g., cathepsins B, H, K, L, and S, and are also known to inactivate cysteine proteinases released by invading microorganisms and parasites. Human cystatins A and B are representatives of family 1 cystatins, also called stefins. Their single nonglycosylated polypeptide chains consist of 98 amino acid residues and have no disulfides. Cystatins A and B have been found primarily in the cytosol of different cell types. Human cystatin C belongs to cystatin family 2. It is a single-chain nonglycosylated protein of 120 residues with two internal disulfide bridges in the C-terminal part of the molecule. Cystatin C has a broad, mainly extracellular, distribution in the organism. Family 1 and 2 cystatins have homologous sequences and share a common fold, comprising a five-stranded antiparallel β-sheet wrapped around a central five-turn α-helix (*6*−*8*).

Cystatins function primarily as emergency inhibitors (*5*) and act by competitively and reversibly forming equimolar complexes with their target proteinases. Inhibition is caused by a wedge-shaped hydrophobic edge of the cystatin molecule being inserted into the proteinase active-site cleft, blocking access of substrates to the active site. This wedge is formed by three structural elements, viz. the N-terminal region, a first binding loop connecting the second and third strands of the β-sheet and a second binding loop between the forth and fifth β-strands (*6*−*8*). These three regions all have residues or consensus motifs conserved in most cystatins, comprising a Gly in the N-terminal region, a Gln-Val-Val-Ala-Gly or similar sequence in the first binding loop and a Leu-Pro or Pro-Trp sequence in the second binding loop of family 1 and 2 cystatins, respectively (Figure 1). The X-ray structure of a complex between cystatin B and a cysteine proteinase, papain (*7*), shows that the three binding regions establish a number of predominantly hydrophobic contacts but also polar interactions with the proteinase. Moreover, site-directed mutagenesis has demonstrated the importance of several individual residues in these regions of different cystatins for the interaction (*9*−*18*).

Computer docking of chicken cystatin, an avian cystatin C analogue, with papain suggests that the enzyme−inhibitor complex can form with no appreciable conformational changes of either protein (*6*). Consistent with this observation, kinetic analyses have shown that cystatins bind to papain and similar proteinases with open active sites in an apparent

† This work was supported by Swedish Research Council-Medicine Grant 4212.

* To whom correspondence should be addressed at the Department of Molecular Biosciences, Section of Veterinary Medical Biochemistry, Swedish University of Agricultural Sciences, Box 575, SE−751 23 Uppsala, Sweden. Tel: +46 18 4714191, Fax: +46 18 550762, E-mail: Ingemar.Bjork@vmk.slu.se.

10.1021/bi030119v CCC: $25.00
Published on Web 09/05/2003

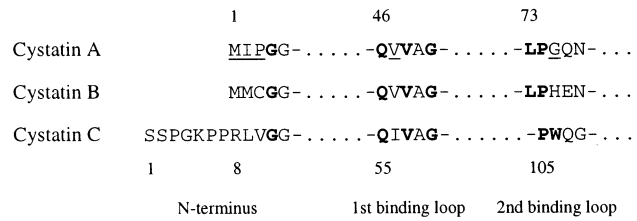

FIGURE 1: Amino acid sequences of the N-terminal regions and the first and second proteinase binding loops of human cystatins A, B, and C. Cystatin A and B numbering is at the top and cystatin C numbering at the bottom. The conserved residues found in most cystatins or, in the case of the second binding loop, in cystatins within either family 1 or 2 are in bold. Cystatin A residues mutated in this work are underlined.

one-step reaction with a rate constant which for many cystatins approaches that of a diffusion-controlled reaction (*19*−*23*). However, cathepsin B, in which a peptide loop partially occludes the active site, is inhibited by cystatins in a two-step reaction. An initial binding of the cystatin N-terminal region to the enzyme in the first step presumably facilitates the occluding loop being displaced in the second step, so that a tight complex can be formed (*24*, *25*). A similar dislocation of the mini-chain of cathepsin H apparently occurs on binding of cystatins to this enzyme (*26*).

Adventitious release of endogeneous cysteine proteinases has been implicated in a number of pathological conditions (*5*, *27*). Cathepsin B is one of the most abundant of such proteinases (*28*), and its efficient inhibition therefore presumably is of considerable pathophysiological importance. Further characterization of the mechanism by which cystatins inhibit cathepsin B should greatly facilitate the design of improved inhibitors of this enzyme. Although cystatin C is the best mammalian cystatin inhibitor of most cysteine proteinases, including papain, cystatins A and C inhibit cathepsin B with comparable affinities, whereas the affinity of cystatin B for this enzyme is appreciably lower (*1*, *3*, *11*, *21*, *23*, *24*, *29*−*32*). The tight binding of cystatin C to cathepsin B is maintained primarily by a high association rate constant, whereas that of cystatin A is affected predominantly by a low dissociation rate constant, the association rate constant being substantially lower than that for cystatin C (*1*, *23*, *24*, *30*, *31*). This lower association rate constant is due to a very slow rate of displacement of the cathepsin B occluding loop by cystatin A in the second binding step (*25*).

The background to these differences in the affinity and kinetics of cathepsin B inhibition by cystatins has been elucidated by site-directed mutagenesis in this work. The three binding regions in cystatin A have been mutated to be identical with or similar to the corresponding regions of cystatin C. Moreover, the second binding loop of cystatin A, which is the binding region differing mostly from the corresponding region of cystatin B, has been altered to be similar to that of the latter inhibitor. The effect of these substitutions on the affinity of cystatin A for cathepsin B, as well as for the model cysteine proteinase, papain, was assessed. We find that grafting of the cystatin C N-terminal region into cystatin A appreciably increases the affinity for cathepsin B, primarily due to an increased association rate constant. This increase was concluded to be due to a considerably increased ability of the mutated cystatin A to displace the occluding loop of cathepsin B in the second binding step. The cystatin A chimera containing the cystatin C N-terminal region was a better inhibitor of cathepsin B than any of the natural cystatins. We further find that introduction of essential features of the second binding loop of either cystatin B or C into cystatin A improves the affinity for papain by favorably affecting both the association and dissociation rate constants.

MATERIALS AND METHODS

Construction of Expression Vectors for Cystatin A Variants. Vectors for expression of N-terminal mutants of human cystatin A were generated by PCR-based mutagenesis of an expression vector for wild-type cystatin A described earlier (*14*, *23*). The original vector carries successive sequences coding for the signal peptide for the outer membrane protein A (OmpA) of *Escherichia coli*, a His-tag[1] and the recognition site for enterokinase, followed by the cystatin A coding sequence. The first three N-terminal residues of cystatin A were replaced by either residues 1−10 or 8−10 of the N-terminal region of human cystatin C (Figure 1). These replacements were introduced by the downstream primers, 5′-CACCAGGCGCGGCGGCTTGCCGGGACTGGACTTGTCGTCGTCGTCGATATGG for N(1−10)CC-cystatin A and 5′-CACCAGGCGCTTGTCGTCGTCGTCGATATGG for N(8−10)CC-cystatin A, which encoded in their 5′-ends the relevant segment of cystatin C (underlined). The upstream primer, 5′-GCTCAGGCGA*CCATGG*GCCATCATCATC, was the same for both variants and contained an *Nco*I cleavage site (in italics). The PCR products were cleaved with *Nco*I, whereas the vector was digested with both *Nco*I and *Sma*I (*14*, *23*). The purified DNA fragments were then cloned into the cleaved vector so that the coding sequences for the cystatin C N-terminal segments were followed by the coding sequence of cystatin A from the codon for Gly4.

Expression vectors for human cystatin A variants with a V47I mutation in the first binding loop or G75H or G75W mutations in the second binding loop (Figure 1) were obtained by a two-step, PCR-based site-directed mutagenesis procedure with the wild-type vector described above as template, essentially as in previous work (*17*, *33*). The mutations were introduced by the mutagenic primers 5′-GAAGCTGTGCAGTATAAAACTCAA**ATT**GTTGCTGGAAC (upstream) and 5′-CTTAATGTAGTAATTTGTTCCAGCAAC**AAT**TTGAGTTTTATAC (downstream) for V47I-cystatin A; 5′-CTTGAAAGTATTCAAAAGTCTTCCC**CAT**CAAAA-

[1] Abbreviations: app, subscript denoting an apparent equilibrium or rate constant determined in the presence of an enzyme substrate; C29A-cathepsin B, human cathepsin B variant in which Cys29 is replaced by Ala; Chaps, 3-[(3-cholamidopropyl)dimethylammonio]-propanesulfonate; DTT, dithiothreitol; G75H- and G75W-cystatin A, human cystatin A variants in which Gly75 is replaced by His and Trp, respectively; H110A/C29A-cathepsin B, C29A-cathepsin B variant in which H110 is replaced by Ala; His-tag, 10 consecutive histidine residues fused to an expressed protein; k_{ass}, bimolecular association rate constant; K_d, dissociation equilibrium constant; k_{diss}, dissociation rate constant; K_i, inhibition constant; k_{obs}, observed pseudo-first-order rate constant; Mes, 4-morpholineethanesulfonic acid; N(1−10)CC- and N(8−10)CC-cystatin A, human cystatin A variants in which the three initial amino-terminal amino acid residues are replaced by residues 1−10 and 8−10 of human cystatin C, respectively; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; V47I-cystatin A, human cystatin A variant in which Val47 is replaced by Ile.

TGAG (upstream) and 5′-GTAAGTACCAAGTCCTCATT-TTG**ATG**GGGAAGAC (downstream) for G75H-cystatin A; 5′-CTTGAAAGTATTCAAAAGTCTTCCC**TGG**CAAAA-TGAG (upstream) and 5′-GTAAGTACCAAGTCCTCATT-TTG**CCA**GGGAAGAC (downstream) for G75W-cystatin A. (Codons for mutated residues are in bold and substitutions in the coding sequence of cystatin A are underlined.) The standard PCR primers, complementary to regions of the plasmid flanking the cystatin A coding sequence, were the same as in ref *17*. The final PCR product, containing the whole coding sequence for the mutant cystatin A, and the wild-type vector were cleaved with *Nco*I and *Bam*HI (*23*), and the purified PCR product was then ligated into the cleaved vector.

Expression and Purification of Cystatin A Variants. The expression vectors for the cystatin A variants were transformed into competent *E. coli*, strain MC1061. The entire cystatin A coding region of plasmids from a number of individual colonies on ampicillin plates were sequenced for identification of plasmids containing the desired mutant sequences.

The recombinant cystatin A variants were expressed in *E. coli* as described earlier (*14*, *23*). The His-tagged proteins were purified from periplasmic extracts (*23*) by immobilized-metal affinity chromatography, either on Ni−NTA Agarose (Qiagen, Hilden, Germany) (*14*) or TALON CellThru (Clontech, Palo Alto, CA). The His-tag was cleaved off by enterokinase (*14*), and the liberated cystatin A mutants were isolated by rechromatography on the same affinity columns. Preparations still contaminated by uncleaved His-tagged proteins were repurified on a TALON Metal Affinity Resin (Clontech) (*17*).

Other Proteins and Determination of Protein Concentration. Papain (EC 3.4.22.2) was purified and stored as inactive *S*-(methylthio)-papain (*34*). The enzyme was >95% homogeneous in SDS−PAGE, had 0.95−1.00 mol of thiol groups/mol of enzyme, and was fully active in binding chicken cystatin (*34*). Human cathepsin B (EC 3.4.22.1) from liver was purchased from Calbiochem (San Diego, CA). An inactive recombinant variant of human cathepsin B, in which the reactive-site Cys29 was replaced by Ala, was a gift from Dr. J. S. Mort (Shriners Hospital for Children, Montreal, Quebec, Canada). This C29A-cathepsin B form was >97% homogeneous in SDS−PAGE and ∼97% active in binding to cystatins (*25*). Chicken cystatin (forms 1 and 2) was isolated from egg white (*34*). Wild-type cystatin A, prepared as in ref *14*, was >99.5% homogeneous in SDS−PAGE and fully active in binding to papain.

Concentrations of all proteins except human liver cathepsin B were calculated from the absorbance at 280 nm with the use of molar absorption coefficients published previously (*14*, *23*, *24*, *34*). The concentration of liver cathepsin B was provided by the manufacturer.

Experimental Conditions. All binding experiments were performed at 25.0 ± 0.2 °C. Papain and liver cathepsin B were activated with 1 mM DTT in the reaction buffer for 10 min before analyses. Interactions with papain were studied in 50 mM Tris-HCl, pH 7.4, containing 100 mM NaCl and 0.1 mM EDTA and, except in stoichiometric titrations and displacement experiments, 1 mM DTT and 0.01% (w/v) Brij 35. The inhibition of active cathepsin B was analyzed in most cases in 50 mM Mes-NaOH, pH 6.0, containing 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, and 0.1% (w/v) poly(ethylene glycol) 6000, and in some cases also in this standard Mes buffer containing in addition 0.1% (w/v) Chaps (*35*). All measurements with active cathepsin B were performed in poly(ethylene glycol) 20000-coated acrylic cuvettes. The buffer in studies of the binding of N(1−10)-CC-cystatin A to C29A-cathepsin B was the standard Mes buffer without DTT (*25*).

Binding Stoichiometry. Stoichiometries of binding of the cystatin A mutants to papain were determined by titrations, monitored by the accompanying decrease of tryptophan fluorescence and performed at least in duplicate, of 1 μM *S*-(methylthio)-papain with the inhibitors (*34*).

Inhibition Constants. K_i for inhibition of cathepsin B by the cystatin A variants was obtained from the equilibrium rates of cleavage of the fluorogenic substrate, carbobenzoxy-L-arginyl-L-arginine 4-methylcoumaryl-7-amide (10 μM; Peptide Institute, Osaka, Japan), by the enzyme at different inhibitor concentrations (*36*, *37*). These concentrations were ≥10-fold higher than the enzyme concentrations and were varied from ∼(1−4) × K_i to ∼(10−20) × K_i.

Kinetics of Inhibition of Active Papain and Cathepsin B. The kinetics of binding of the cystatin A variants to the two proteinases were studied under pseudo-first-order conditions in the presence of a fluorogenic substrate by continuously monitoring the fluorescence increase (*37*). The substrate for papain was 10 μM carbobenzoxy-L-phenylalanyl-L-arginine 4 methylcoumaryl-7-amide (Peptide Institute), whereas that for cathepsin B and its concentration were the same as in the K_i determinations. Concentrations of cystatin A variants were ≥10 higher than enzyme concentrations and were varied in a 10−20-fold range, up to ∼2−8 nM and ∼60−600 nM for the inhibition of papain and cathepsin B, respectively.

Stopped-Flow Kinetics. The rate of association of N(1−10)CC-cystatin A with inactive C29A-cathepsin B was analyzed by continuously monitoring the increase in intrinsic tryptophan fluorescence accompanying complex formation in a stopped-flow fluorometer (SX-17MV; Applied Biophysics, Leatherhead, UK) (*24*, *25*). Inhibitor concentrations were 10-fold higher than C29A-cathepsin B concentrations, ensuring pseudo-first-order conditions.

Dissociation Kinetics. k_{diss} for complexes between the cystatin A mutants and papain was determined by displacement experiments, in which a large excess of chicken cystatin was used to essentially irreversibly trap papain dissociating from the complexes (*19*, *21*). The initial complex concentrations were 2.5−4.0 μM, with a molar ratio of the inhibitors to papain of 1.1, and the molar ratio of chicken cystatin (form 2) to cystatin A mutants was 10−30. The slow dissociation was followed for 100−150 h by measuring the increase in the concentration of the newly formed chicken cystatin−papain complex by ion-exchange chromatography on a 1 mL MonoQ column (Amersham Biosciences, Uppsala, Sweden).

Miscellaneous Procedures. N-terminal sequences were analyzed in an Applied Biosystems 477A Protein Sequencer. Molecular masses were measured in a Kratos Kompact MALDI 4 mass spectrometer (Kratos Analytical, Manchester, UK) (*17*, *23*). SDS−PAGE under reducing and nonreducing conditions was done with the Tricine buffer system (*38*).

Modeling. A model of the complex between human wild-type cystatin A and papain, developed previously from the

Table 1: Dissociation Equilibrium Constants and Association and Dissociation Rate Constants for the Interaction of Recombinant Human Wild-Type Cystatins and Cystatin A Variants with Papain and Human Cathepsin B at 25 °C[a]

proteinase	cystatin form	K_d (M)	k_{ass} (M^{-1} s^{-1})	k_{diss} (s^{-1})
papain	cystatin A	1.8×10^{-13}	3.1×10^{6}	5.5×10^{-7}
	cystatin B	4.9×10^{-14}	9.7×10^{6}	4.8×10^{-7}
	cystatin C	1.1×10^{-14}	1.1×10^{7}	1.3×10^{-7}
	N(1−10)CC-cystatin A	1.3×10^{-13} [b]	$(2.74 \pm 0.05) \times 10^{6}$ (9)	$(3.6 \pm 0.3) \times 10^{-7}$ (3)
		[0.7]	[0.9]	[0.7]
	N(8−10)CC-cystatin A	9.8×10^{-14} [b]	$(3.05 \pm 0.02) \times 10^{6}$ (8)	$(3.0 \pm 0.1) \times 10^{-7}$ (3)
		[0.5]	[1]	[0.5]
	V47I-cystatin A	3.6×10^{-13} [b]	$(3.26 \pm 0.02) \times 10^{6}$ (10)	$(1.17 \pm 0.04) \times 10^{-6}$ (3)
		[2]	[1.1]	[2.1]
	G75H-cystatin A	1.6×10^{-14} [b]	$(6.9 \pm 0.2) \times 10^{6}$ (8)	$(1.1 \pm 0.1) \times 10^{-7}$ (3)
		[0.1]	[2.2]	[0.2]
	G75W-cystatin A	1.4×10^{-14} [b]	$(1.41 \pm 0.02) \times 10^{7}$ (8)	$(2.0 \pm 0.1) \times 10^{-7}$ (3)
		[0.1]	[4.5]	[0.4]
cathepsin B	cystatin A	9.1×10^{-10}	3.9×10^{4}	3.5×10^{-5}
		$(1.1 \pm 0.1) \times 10^{-9}$ (11)[c]	$(4.6 \pm 0.2) \times 10^{4}$ (8)[c]	5.1×10^{-5} [c,d]
	cystatin B	1.8×10^{-8}	3.0×10^{5}	5.4×10^{-3}
	cystatin C	2.8×10^{-10}	2.4×10^{6}	6.7×10^{-4}
	N(1−10)CC-cystatin A	$\leq 9.6 \times 10^{-11}$ (3)	$(3.70 \pm 0.05) \times 10^{5}$ (12)	$\leq 3.8 \times 10^{-5}$ [d]
		[≤0.1]	[9]	[≤1.1]
		$(6.9 \pm 0.4) \times 10^{-11}$ (9)[c]	$(2.88 \pm 0.07) \times 10^{5}$ (5)[c]	2×10^{-5} [c,d]
		[0.06]	[6]	[0.4]
	N(8−10)CC-cystatin A	$\leq 2.4 \times 10^{-10}$ (4)	$(1.77 \pm 0.05) \times 10^{5}$ (8)	$\leq 4.2 \times 10^{-5}$ [d]
		[≤0.3]	[4.5]	[≤1.2]
	V47I-cystatin A	$(4.8 \pm 0.3) \times 10^{-9}$ (7)	$(3.59 \pm 0.05) \times 10^{4}$ (8)	1.7×10^{-4} [d]
		[5.3]	[0.9]	[4.9]
	G75H-cystatin A	$(1.55 \pm 0.05) \times 10^{-9}$ (6)	$(2.98 \pm 0.07) \times 10^{4}$ (10)	4.7×10^{-5} [d]
		[1.7]	[0.8]	[1.3]
	G75W-cystatin A	$(3.0 \pm 0.2) \times 10^{-9}$ (7)	$(7.0 \pm 0.2) \times 10^{4}$ (8)	2.1×10^{-4} [d]
		[3.3]	[1.8]	[6]

[a] Experimental conditions are described in Materials and Methods. Analyses with cathepsin B were done in the standard Mes buffer, except where indicated. Values determined experimentally in this work are given with the SEM and with the number of experiments in parentheses. Calculated values and values for wild-type cystatins, taken from previous work (*21*, *23*, *24*, *32*) and shown for comparison, are given without errors. Numbers in square brackets are relative values, calculated as the ratio of the corresponding constant to that for wild-type cystatin A, measured under the same conditions. [b] Calculated from k_{ass} and k_{diss}. [c] Measured in the standard Mes buffer containing 0.1% Chaps. [d] Calculated from K_i and k_{ass}.

X-ray structure of the human C3S-cystatin B−*S*-(carboxymethyl)−papain complex (*7*, *17*), was used as a basis for further modeling. This original model was derived by a procedure involving optimizing the conformation of the second binding loop of cystatin A in the complex. The model was further refined by energy minimization with the program Swiss-PdbViewer (url: http://www.expasy.ch/spdbv/) (*39*). Structures of the complexes of G75H- and G75W-cystatin A with papain were then derived by replacing Gly75 of cystatin A in the model by His or Trp by the mutation facility of the program. The two models so obtained were corrected by the program facility "Quick and Dirty Fixing" of all side chains in each complex and were refined by the program facility "Exhaustive Search Fixing" and energy minimization.

RESULTS

Characterization of Recombinant Cystatin A Mutants. The background to the differences in cathepsin B inhibition by different cystatins was investigated by construction of recombinant cystatin A variants with mutations in the three polypeptide chain regions important for inhibition of target proteinases. In two N-terminal mutants, N(1−10)CC- and N(8−10)CC-cystatin A, the first three N-terminal residues of cystatin A were replaced by all 10 or only the three proximal residues, respectively, preceding the conserved Gly11 in the N-terminal region of cystatin C (Figure 1). In three other mutants, single amino acid replacements were introduced in either the first or second proteinase binding loops. The V47I mutation in the first binding loop made the latter identical to the corresponding loop of cystatin C, while the G75H and G75W mutations in the second binding loop made this loop similar to the corresponding loop of cystatins B and C, respectively (Figure 1).

The His-tagged recombinant variants were expressed in *E. coli*, and the His-tag was removed by enterokinase. Cleavage of His-tagged cystatin A variants with an unmodified N-terminal region, like the V47I, G75H, and G75W mutants studied in this work, by enterokinase has been shown previously to give an authentic N-terminus (*14*, *17*). Sequencing of five residues of the N(1−10)CC and N(8−10)CC mutants confirmed proper cleavage also of these variants and was consistent with the presence of the desired substitutions in the N-terminal region. The molecular masses of all mutants, measured by MALDI mass spectrometry, corresponded within 6 Da to the expected values, indicating that the mutants had the correct lengths and sequences. All cystatin A variants were >99.5% homogeneous in SDS−PAGE and had a stoichiometry of binding to *S*-(methylthio)-papain of 0.95−0.98, i.e., were essentially fully active in binding of cysteine proteinases.

Binding Affinity. The affinity and kinetics of interaction of the cystatin A variants with papain and cathepsin B were characterized. The high affinity of all variants for papain precluded a direct determination of K_d. Therefore, K_d for these interactions was calculated from k_{ass} and k_{diss} measured in separate kinetic experiments (Table 1; see also below).

None of the substitutions at the N-terminus, nor the V47I mutation in the first binding loop of cystatin A, influenced the affinity for papain to any essential extent. Importantly, however, the G75H and G75W mutations in the second binding loop led to a ~10-fold higher affinity for papain, compared with that of wild-type cystatin A.

K_d for the complexes between the cystatin A mutants and cathepsin B were determined as K_i from the equilibrium rates of cleavage of a fluorogenic substrate by the proteinase at varying inhibitor concentrations (Table 1). Analyses with N(1−10)CC- and N(8−10)CC-cystatin A in the conventional Mes buffer used in previous work showed that both these variants had a higher affinity for cathepsin B than the wild-type inhibitor. However, as a consequence, only upper limits of K_d could be estimated under these conditions, as cathepsin B was not sufficiently stable during the long reaction times necessary for accurate determinations. These upper limits of K_i were >10- and >3-fold lower for the N(1−10)CC and N(8−10)CC mutants, respectively, than K_i for the wild-type inhibitor. Addition of 0.1% (w/v) of the zwitterionic detergent Chaps (*35*) to the standard Mes buffer was found to considerably, for up to 7 h, stabilize cathepsin B, so that higher affinities could be measured. K_i for the binding of wild-type cystatin A to cathepsin B was similar in the presence and absence of Chaps, showing that this detergent negligibly influences the affinity of the interaction. K_i for N(1−10)CC-cystatin A measured under these conditions was ~16-fold lower than that for wild-type cystatin A, revealing a considerably higher affinity of the N(1−10)CC mutant than of the wild-type inhibitor for cathepsin B. In contrast to the N-terminal mutations, the mutations within the first and second binding loops resulted in moderately, 2−5-fold, reduced affinities of cystatin A for cathepsin B that could be quantified in the standard Mes buffer.

Association Kinetics. k_{ass} for the binding of the cystatin A variants to papain and cathepsin B was determined from the kinetics of inhibition of the activities of the enzymes against fluorogenic substrates. The resulting progress curves were well fitted by a single-exponential function (*37*). $k_{obs,app}$ derived from these fits varied linearly with the inhibitor concentration within the range covered, and k_{ass} was obtained from the slopes of these plots. k_{diss} could not be determined from the plots, as the intercepts on the ordinate were undistinguishable from zero.

The variants with mutations in the N-terminal region or the first binding loop had k_{ass} values for papain binding essentially identical to that of wild-type cystatin A (Table 1). However, both mutations in the second binding loop, in particular, G75W, resulted in an increased rate of association with the enzyme. The G75H and G75W mutants thus had ~2-fold and almost 5-fold higher k_{ass}, respectively, than the wild-type inhibitor.

Contrary to what was observed for papain, k_{ass} for N(1−10)CC- and N(8−10)CC-cystatin A binding to cathepsin B was ~9- and ~5-fold higher, respectively, than k_{ass} for the wild-type inhibitor (Table 1). However, the mutations within the first and the second binding loops had no or only a minor (<2-fold) effect on k_{ass} for inhibition of this enzyme. Analyses in the presence of 0.1% (w/v) Chaps showed that the kinetics of binding of wild-type and N(1−10)CC-cystatin A to cathepsin B were not perturbed by this detergent.

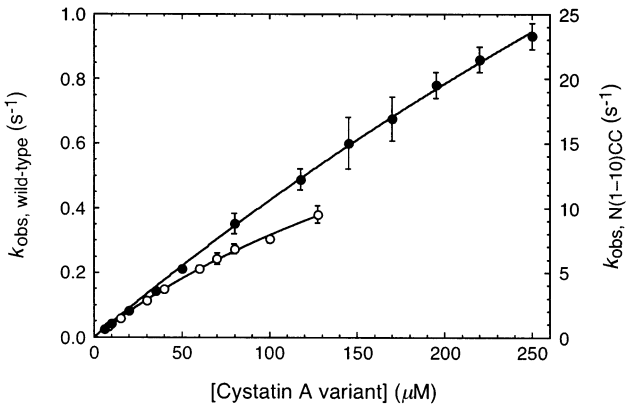

FIGURE 2: Observed pseudo-first-order rate constants, k_{obs}, for the binding of wild-type and N(1−10)CC-cystatin A to C29A-cathepsin B as a function of inhibitor concentration. (○) Wild-type cystatin A; (●) N(1−10)CC-cystatin A. Note the different scales of the two plots. The data for wild-type cystatin A are taken from earlier work (*25*) and are shown for comparison. Values are averages ± SEM of 10−20 measurements; errors bars not seen lie within the dimensions of the symbols. The solid lines are nonlinear least-squares regression fits of the data to the hyperbolic function characterizing the two-step binding mechanism (*24*, *25*).

Table 2: Kinetic Constants for the Two-Step Mechanism of Binding of Recombinant Human Cystatins A and C and N(1−10)CC-Cystatin A to Recombinant Human C29A-Cathepsin B at pH 6.0, 25 °C[a]

cystatin form	K_1 (μM)	k_{+2} (s^{-1})	k_{ass} ($M^{-1}\ s^{-1}$)
wild-type cystatin A	290 ± 40	1.2 ± 0.1	4.2×10^3
N(1−10)CC-cystatin A	1200 ± 200	130 ± 20	1.1×10^5
wild-type cystatin C	70 ± 10	90 ± 10	1.3×10^6

[a] The dissociation equilibrium constant of the first step, K_1, and the forward rate constant of the second step, k_{+2}, of the two-step binding mechanism (*24*, *25*) were obtained by nonlinear regression analyses of the data in Figure 2 or of similar data for wild-type cystatin C. The overall association rate constant, k_{ass}, was calculated as k_{+2}/K_1. The buffer was the standard Mes buffer without DTT. The values for wild-type cystatins are taken from previous work (*25*) and are shown for comparison. Measured values are given with the SEM, whereas calculated values are given without errors.

The background to the increased k_{ass} for the interaction of the N-terminal cystatin A mutants with cathepsin B was elucidated by stopped-flow analyses of the binding of N(1−10)-CC-cystatin A to an inactive C29A-variant of cathepsin B at high inhibitor concentrations. The binding was monitored under pseudo-first-order conditions by the fluorescence change accompanying the interaction (*25*). All fluorescence traces were well fitted by a single-exponential function, giving k_{obs}. As in similar previous studies of the binding of wild-type cystatins A and C to wild-type and C29A-cathepsin B, k_{obs} showed a hyperbolic dependence on inhibitor concentration (Figure 2), indicative of a two-step binding mechanism (*24*, *25*). In this mechanism, a weak bimolecular complex is initially formed in a rapid equilibrium, followed by a conformational change that has been shown to involve displacement of the occluding loop of cathepsin B by the inhibitor and lead to formation of a stable cystatin−proteinase complex (*25*). The N(1−10)CC substitution resulted in only a small, ~4-fold, increase in the dissociation equilibrium constant for the first step of this reaction, K_1, compared with the value for wild-type cystatin A (Table 2). In contrast, the forward rate constant of the second step, k_{+2}, was increased

a substantial ~100-fold and was similar to that for the interaction between cystatin C and C29A-cathepsin B. As a consequence of the changes in K_1 and k_{+2}, the value of k_{ass}, obtained from the initial slopes of the plots in Figure 2, was ~30-fold higher for N(1−10)CC-cystatin A binding than for wild-type cystatin A binding to C29A-cathepsin B (Table 2).

Dissociation Kinetics. The low values of k_{diss} for the complexes between the cystatin A mutants and papain were determined by displacement experiments similar to those used previously in studies of the interactions of wild-type cystatins A, B, and C with this enzyme (*21*, *23*, *32*). Papain dissociating from the complexes with the cystatin A mutants was essentially irreversibly trapped by an excess of chicken cystatin, and the progress of the dissociation was monitored by analyses by ion-exchange chromatography of the amount of the new complexes formed between papain and chicken cystatin after different incubation times. Most mutations marginally affected the stability of the complexes with papain, whereas the G75H and G75W mutations in the second binding loop led to moderate, ~3−5-fold, decreases in k_{diss} (Table 1).

The large amounts of enzyme necessary for displacement experiments precluded direct measurements of k_{diss} for the complexes of the cystatin A mutants with cathepsin B. Values of k_{diss} for these interactions were therefore calculated from experimentally measured values of K_d and k_{ass}. The substitutions in the N-terminal region and the G75H mutation in the second binding loop minimally affected the stability of the complex with cathepsin B (Table 1). However, the V47I mutation in the first binding loop and the G75W mutation in the second binding loop moderately decreased this stability by increasing k_{diss} ~5-fold.

DISCUSSION

A major result of this work is that replacement of the N-terminal residues of cystatin A preceding the conserved Gly, MIP, by the corresponding cystatin C N-terminal region, SSPGKPPRLV (Figure 1), improves the affinity of cystatin A for cathepsin B by ~15-fold. This increase is equivalent to the N-terminal region of cystatin C contributing ~7 kJ mol^{-1} more to the energy of binding to cathepsin B than the authentic segment of wild-type cystatin A. The improved affinity was predominantly due to an increased k_{ass}. This rate enhancement presumably was a consequence of an increased forward rate constant of the conformational change in the second step of the two-step mechanism of cathepsin B binding (*24*, *25*). Although evidence for such an increase was obtained with an inactive cathepsin B variant, in which the reactive-site Cys29 was replaced by Ala, previous work has shown that this substitution does not appreciably influence the kinetics of cathepsin B binding by cystatins (*25*). The results therefore strongly indicate that the forward rate constant of the conformational change on binding of the N-terminally substituted cystatin A to wild-type cathepsin B is increased to a comparable extent, ~100-fold, as that on binding to the inactive cathepsin B variant. The transplanted cystatin C N-terminal region in the chimeric inhibitor thus apparently promotes displacement of the occluding loop of cathepsin B in the second binding step appreciably better than the corresponding region in wild-type cystatin A (*25*). This augmenting effect may be due to a more productive binding of the cystatin C N-terminal region in the first binding step, inducing an orientation of the first and second binding loops of the chimera that is more advantageous for displacement of the occluding loop. A similar favorable binding of the N-terminal region presumably is mainly responsible for the high rate of association of cystatin C with cathepsin B. Arg8 in the cystatin C N-terminal region may be particularly important for inducing this optimal binding mode, although Leu9 and Val10 also contribute (*11*, *13*). However, the entire N-terminal region of cystatin C introduced into cystatin A increased the rate of dislocation of the cathepsin B occluding loop more efficiently than only the RLV segment of this region (Figure 1), as judged by the higher k_{ass} attained. Interactions established by residues preceding the RLV segment therefore also appear to be important for optimal productive binding of the N-terminal region. The full rate of cathepsin B binding by cystatin C was not attained even by grafting of the entire N-terminal region of this inhibitor into cystatin A, perhaps due to residues C-terminal of the conserved Gly11 in cystatin C being of importance for the higher association rate. These residues may mediate an efficient initial binding of cystatin C to cathepsin B either by interacting directly with the enzyme or promoting the contacts made by the N-terminal region. Notably, the considerably, ~30-fold, lower k_{diss} for the binding of cystatin A than of cystatin C to cathepsin B was retained by the cystatin A variant containing the entire cystatin C N-terminal region. As a consequence, this chimeric inhibitor had ~4-fold higher affinity for cathepsin B than wild-type cystatin C and therefore a higher affinity for this enzyme than any natural cystatin characterized (*3*, *11*, *24*, *37*).

In contrast to the effect on cathepsin B inhibition, grafting of the RLV segment or the entire cystatin C N-terminal region into cystatin A minimally affected the affinity and kinetics of binding of the inhibitor to papain. The RLV segment thus apparently interacts with papain in a similar manner as the authentic MIP segment of cystatin A, contributing ~40% of the total unitary free energy of binding by maintaining a low k_{diss} (*15*). This conclusion is in agreement with results of a previous study by random mutagenesis of the first four N-terminal residues of cystatin A and phage-display selection of high-affinity variants, showing that several, mainly hydrophobic, sequences can substitute for the MIP fragment of cystatin A in tight binding to papain (*40*). The lack of an effect of introducing the entire cystatin C N-terminal region into cystatin A on papain binding further indicates that residues N-terminal of the RLV segment do not enhance the interaction with this enzyme.

Previous studies have shown that replacement of the N-terminal region of human cystatin B with parts of the corresponding regions of cystatin C or kininogen does not affect the affinity of this inhibitor for papain or cathepsin L (*41*). This finding is in agreement with the results for papain inhibition by the similarly substituted cystatin A in our work. However, these replacements were found to decrease the affinity of cystatin B for cathepsin B (*41*), in contrast with the increase in affinity observed for cathepsin B inhibition by the N-terminally substituted cystatin A in this work. These opposite effects of transplanting the cystatin C N-terminal

region into the two inhibitors may be related to the fact that cystatin B is an appreciably weaker inhibitor of cathepsin B than cystatin A (*1*, *3*, *23*, *32*). The substitutions of cystatin B in the earlier studies also decreased the affinity for cathepsin H (*41*), perhaps because the N-terminal regions introduced interfered with the accommodation of the mini-chain of cathepsin H in the complex (*26*).

Replacement of Val47 in the first binding loop of cystatin A with Ile in this work, making the loop identical to that of cystatin C, decreased the affinity for cathepsin B by ~5-fold, exclusively due to an increase in k_{diss}. This rate constant thereby approached that of the binding of cystatin C to cathepsin B. The bulkier Ile thus adversely affects the fit of the first binding loop of cystatins into the cathepsin B active-site cleft. However, the effect of the V47I mutation on papain binding was negligible, indicating that Val and Ile in position 47 in the first binding loop are essentially equivalent in cystatin binding to this enzyme.

Substitution of Gly75 in the second binding loop of cystatin A by His, which makes this loop similar to that of cystatin B, did not influence the affinity or kinetics of binding of the inhibitor to cathepsin B. However, substitution by Trp, to simulate the essential Pro-Trp binding motif of family 2 cystatins, reduced the affinity ~5-fold, mainly due to an increased k_{diss}. Hence, a His in position 75 can be well accommodated in the complex of cystatin A with cathepsin B, whereas the bulky Trp perturbs somewhat the intermolecular contacts stabilizing the complex. Interestingly, however, both the His and Trp substitutions increased the affinity of cystatin A for papain appreciably, ~10-fold. The increase caused by the G75H mutation was predominantly due to a stabilization of the complex with papain by a decreased k_{diss}. In contrast, the enhanced affinity resulting from the G75W-cystatin A mutation was mainly due to a faster association rate, k_{ass} being increased up to that for cystatin C binding to papain (*21*) and approaching the value expected for a diffusion-limited rate. As a consequence, the affinities of both mutated cystatin A variants for papain were similar to that of cystatin C for this enzyme, the highest affinity for papain known and close to the highest affinity for a protein−protein interaction demonstrated (*3*, *21*, *42*). These changes could have been due to the mutations having induced local conformational alterations of the second binding loop, perhaps decreasing the mobility of the loop and thereby promoting both an increased association rate and new interactions leading to a tighter binding of the loop into the active-site cleft of papain. The increased affinity for papain caused by introduction of His or Trp into the second binding loop are in agreement with previous studies showing that the two residues are essential for the tight binding of cystatins B and C, respectively, to this enzyme (*12*, *16*).

Interactions between the His or Trp replacing Gly75 in cystatin A and papain that could contribute to an increased stability of the complexes were indicated by computer modeling. In the models, which were derived from the coordinates for the cystatin B−papain complex (*7*), both His and Trp could be accommodated in position 75 without steric hindrance. The model of the complex between the His variant and papain revealed the possibility of a hydrogen bond between the His side chain and the main chain carbonyl group of Gly180 of papain. Moreover, the model of the complex of the Trp variant with the enzyme indicated two close hydrophobic contacts (~3.8 Å) between the Trp residue and Gly180 of papain. These interactions, which were absent in the model of the complex with wild-type cystatin A, might be responsible for the increased stability of the complexes of the mutant inhibitors with papain that was observed experimentally. In both models, however, certain hydrophobic contacts of Leu73 and Pro74 in the second binding loop with papain were somewhat longer than in the model of the complex with wild-type cystatin A, which might be expected to result in a weakened binding. Nevertheless, this attenuation might be more than compensated by the new interactions indicated by the models.

In conclusion, the results show that essential binding properties of the most potent mammalian inhibitor of papain-like cysteine proteinases, cystatin C, can be bestowed on the weaker inhibitor, cystatin A, by making either the N-terminal region or the second binding loop of cystatin A similar or identical to the corresponding region of cystatin C. Grafting of the N-terminal region of cystatin C into cystatin A creates a fast and efficient inhibitor of cathepsin B that is more potent against this proteinase than any natural inhibitor isolated so far. Moreover, introduction of a cystatin C-like, and also of a cystatin B-like, sequence into the second binding loop of cystatin A substantially improves papain inhibition by this inhibitor. These engineered enhancements of the inhibitory ability of cystatins indicate possible pathways for the design of new and improved inhibitors of lysosomal cysteine proteinases for potential medical applications.

ACKNOWLEDGMENT

We are grateful to Dr. John S. Mort (Joint Diseases Laboratory, Shriners Hospital for Children and the Department of Surgery, McGill University, Montreal, Quebec, Canada) for the kind gift of recombinant human C29A-cathepsin B. We also thank Dr. Åke Engström (Department of Medical Biochemistry and Microbiology, Uppsala University) for molecular mass determinations and amino acid sequencing.

REFERENCES

1. Barrett, A. J., Rawlings, N. D., Davies, M. E., Machleidt, W., Salvesen, G., and Turk, V. (1986) in *Proteinase Inhibitors* (Barrett, A. J. and Salvesen, G., Eds.) pp 515−569, Elsevier, Amsterdam.
2. Turk, V., and Bode, W. (1991) *FEBS Lett. 285*, 213-219.
3. Abrahamson, M. (1994) *Methods Enzymol. 244*, 685−700.
4. Turk, B., Turk, V., and Turk, D. (1997) *Biol. Chem. 378*, 141−150.
5. Turk, B., Turk, D., and Salvesen, G. S. (2002) *Curr. Pharm. Des. 8*, 1623−1637.
6. 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.
7. Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B., and Turk, V. (1990) *EMBO J. 9*, 1939−1947.
8. Martin, J. R., Craven, C. J., Jerala, R., Kroon-Zitko, L., Zerovnik, E., Turk, V., and Waltho, J. P. (1995) *J. Mol. Biol. 246*, 331−343.
9. Auerswald, E. A., Genenger, G., Assfalg-Machleidt, I., Machleidt, W., Engh, R. A., and Fritz, H. (1992) *Eur. J. Biochem. 209*, 837−845.
10. Björk, I., Brieditis, I., and Abrahamson, M. (1995) *Biochem. J. 306*, 513−518.
11. Hall, A., Håkansson, K., Mason, R. W., Grubb, A., and Abrahamson, M. (1995) *J. Biol. Chem. 270*, 5115−5121.
12. Björk, I., Brieditis, I., Raub-Segall, E., Pol, E., Håkansson, K., and Abrahamson, M. (1996) *Biochemistry 35*, 10720−10726.

13. Mason, R. W., Sol-Church, K., and Abrahamson, M. (1998) *Biochem. J. 330*, 833−838.
14. Estrada, S., Nycander, M., Hill, N. J., Craven, C. J., Waltho, J. P., and Björk, I. (1998) *Biochemistry 37*, 7551−7560.
15. Estrada, S., Pavlova, A., and Björk, I. (1999) *Biochemistry 38*, 7339−7345.
16. Pol, E., and Björk, I. (1999) *Biochemistry 38*, 10519−10526.
17. Pavlova, A., Estrada, S., and Björk, I. (2002) *Eur. J. Biochem. 269*, 5649−5658.
18. Pol, E., and Björk, I. (2003) *Biochim. Biophys. Acta 1645*, 105−112.
19. Björk, I., Alriksson, E., and Ylinenjärvi, K. (1989) *Biochemistry 28*, 1568−1573.
20. Björk, I., and Ylinenjärvi, K. (1990) *Biochemistry 29*, 1770−1776.
21. Lindahl, P., Abrahamson, M., and Björk, I. (1992) *Biochem. J. 281*, 49−55.
22. Turk, B., Colic, A., Stoka, V., and Turk, V. (1994) *FEBS Lett. 339*, 155−159.
23. Pol, E., Olsson, S. L., Estrada, S., Prasthofer, T. W., and Björk, I. (1995) *Biochem. J. 311*, 275−282.
24. Nycander, M., Estrada, S., Mort, J. S., Abrahamson, M., and Björk, I. (1998) *FEBS Lett. 422*, 61−64.
25. Pavlova, A., Krupa, J. C., Mort, J. S., Abrahamson, M., and Björk, I. (2000) *FEBS Lett. 487*, 156−160.
26. Jenko, S., Dolenc, I., Guncar, G., Dobersek, A., Podobnik, M., and Turk, D. (2003) *J. Mol. Biol. 326*, 875−885.
27. Turk, B., Turk, D., and Turk, V. (2000) *Biochim. Biophys. Acta 1477*, 98−111.
28. Xing, R. Y., Addington, A. K., and Mason, R. W. (1998) *Biochem. J. 332*, 499−505.
29. Abrahamson, M., Barrett, A. J., Salvesen, G., and Grubb, A. (1986) *J. Biol. Chem. 261*, 11282−11289.
30. Abrahamson, M., Mason, R. W., Hansson, H., Buttle, D. J., Grubb, A., and Ohlsson, K. (1991) *Biochem. J. 273*, 621−626.
31. Turk, B., Ritonja, A., Björk, I., Stoka, V., Dolenc, I., and Turk, V. (1995) *FEBS Lett. 360*, 101−105.
32. Pol, E., and Björk, I. (2001) *Protein Sci. 10*, 1729−1738.
33. Higuchi, R., Krummel, B., and Saiki, R. K. (1988) *Nucleic Acids Res. 16*, 7351−7367.
34. Lindahl, P., Alriksson, E., Jörnvall, H., and Björk, I. (1988) *Biochemistry 27*, 5074−5082.
35. Krupa, J. C., and Mort, J. S. (2000) *Anal. Biochem. 283*, 99−103.
36. Lindahl, P., Nycander, M., Ylinenjärvi, K., Pol, E., and Björk, I. (1992) *Biochem. J. 286*, 165−171.
37. Björk, I., Pol, E., Raub-Segall, E., Abrahamson, M., Rowan, A. D., and Mort, J. S. (1994) *Biochem. J. 299*, 219−225.
38. Schägger, H., and von Jagow, G. (1987) *Anal. Biochem. 166*, 368−379.
39. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis 18*, 2714−2723.
40. Ylinenjärvi, K., Widersten, M., and Björk, I. (1999) *Eur. J. Biochem. 261*, 682−688.
41. Jerala, R., Kroon-Zitko, L., Popovic, T., and Turk, V. (1994) *Eur. J. Biochem. 224*, 797−802.
42. Vincent, J.-P., and Lazdunski, M. (1972) *Biochemistry 11*, 2967−2977.

BI030119V