# The Importance of the Second Hairpin Loop of Cystatin C for Proteinase Binding. Characterization of the Interaction of Trp-106 Variants of the Inhibitor with Cysteine Proteinases†

Ingemar Björk,*,‡ Ingrid Brieditis,‡ Elke Raub-Segall,‡ Ewa Pol,‡ Katarina Håkansson,§ and Magnus Abrahamson§

*Department of Veterinary Medical Chemistry, Swedish University of Agricultural Sciences, Uppsala Biomedical Center, Box 575, S-751 23 Uppsala, Sweden, and Department of Clinical Chemistry, University of Lund, University Hospital, S-221 85 Lund, Sweden*

*Received February 21, 1996; Revised Manuscript Received May 15, 1996*⊗

ABSTRACT: The single Trp of human cystatin C, Trp-106, is located in the second hairpin loop of the proteinase binding surface. Substitution of this residue by Gly markedly altered the spectroscopic changes accompanying papain binding and reduced the affinity for papain, actinidin, and cathepsins B and H by 300−900-fold. The decrease in affinity indicated that the side chain of Trp-106 contributes a similar free energy, −14 to −17 kJ·mol$^{-1}$, to the binding to all four cysteine proteinases, corresponding to about 20−30% of the total binding energy. Replacement of Trp-106 by Phe led to a smaller (30−120-fold) decrease in affinity for the four enzymes than Gly substitution. The binding energy of the Phe residue corresponded to 20−45% of that of Trp, showing that a phenyl group can only partly substitute for the indole ring. The reduced affinities of the cystatin C Trp-106 variants for all proteinases studied were due almost exclusively to increased dissociation rate constants. The second hairpin loop thus contributes to the binding primarily by keeping cystatin C anchored to the proteinase once the complex has been formed. This role is partly in contrast to that of the N-terminal region, which increases the affinity of cystatin C for cathepsin B by increasing the association rate constant. Removal of the N-terminal region of the Trp-106→Gly variant by proteolytic cleavage substantially weakened the binding to papain and cathepsin B. The resulting affinity indicated that the first hairpin loop (the "QVVAG-region"), which is the only region of the proteinase binding surface remaining intact in the truncated variant, contributes 40−60% of the total free energy of binding of cystatin C to both proteinases.

Cystatin C is the predominant mammalian low-molecular-mass extracellular inhibitor of cysteine proteinases. Together with other members of the cystatin superfamily of cysteine proteinase inhibitors, it serves to protect the organism against undesirable proteolysis by endogenous or exogenous enzymes. Cystatins inactivate their target proteinases by trapping them in a reversible, tight equimolar complex. The crystal structures of chicken cystatin and of a complex between cystatin B (also called stefin B) and papain have provided considerable information on the nature of these interactions (Bode et al., 1988; Stubbs et al., 1990). The structures show that the inhibitors have a wedge-shaped binding surface composed of three segments of the polypeptide chain, viz., a flexible amino-terminal region, a first hairpin loop in the middle of the chain, and a second hairpin loop close to the carboxy-terminal end. The crystal structures, together with modeling studies, further indicate that the tight binding of target proteinases is due to this wedge interacting in a highly complementary fashion with the active-site cleft of the enzymes.

The contributions of the amino-terminal region and the first hairpin loop of cystatin C or the homologous inhibitor, chicken cystatin, to the interaction with target proteinases have been characterized in detail by proteolytic truncation and by deletion and site-directed mutagenesis (Abrahamson et al., 1987, 1991; Machleidt et al., 1989; Auerswald et al., 1992, 1994, 1995; Hall et al., 1993, 1995; Björk et al., 1994, 1995). The proposal that the second hairpin loop participates in the binding has been verified by chemical modification and deletion mutagenesis of chicken cystatin (Nycander & Björk, 1990; Auerswald et al., 1995). Moreover, initial studies by site-directed mutagenesis of cystatin C have indicated that Trp-106 of this loop is involved in the binding (Hall et al., 1995). In this work, we have characterized the contribution of Trp-106 of cystatin C to the inhibitory activity by spectroscopic, equilibrium, and kinetic studies of the interactions between a series of cysteine proteinases and cystatin C variants having this residue replaced by Phe or Gly.

## MATERIALS AND METHODS

*Proteins.* Cystatin C variants in which Trp-106 was replaced with Phe or Gly (W106F- and W106G-cystatin C,[1] respectively) were produced by *Escherichia coli* expression. The appropriate genes were created by a general mutagenesis protocol detailed earlier (Hall et al., 1995). Plasmid pHD313 (Abrahamson et al., 1988), containing a full-length cDNA encoding human cystatin C and modified to include the *E. coli* outer membrane protein A signal peptide to direct the

† This work was supported by grants from the Swedish Medical Research Council (4212 to I. Björk and 9915 to M.A.), the Medical Faculty of the University of Lund (to M.A.), and the Magnus Bergvall's Foundation (to M.A.).

* To whom correspondence should be addressed at the Department of Veterinary Medical Chemistry, Swedish University of Agricultural Sciences, Box 575, S-751 23 Uppsala, Sweden. Telephone: +46 18 174191. Fax: +46 18 550762. E-mail: Ingemar.Bjork@vmk.slu.se.

‡ Swedish University of Agricultural Sciences.

§ University of Lund.

⊗ Abstract published in *Advance ACS Abstracts,* July 15, 1996.

S0006-2960(96)00420-5 CCC: $12.00 

recombinant protein to the bacterial periplasm, was the starting material. Unique recognition sites in pHD313 for *Bgl*II and *Eco*RI were used to excise a 357 bp fragment that included the coding sequence for cystatin C residues 101−120 and the 3′ noncoding region. In two separate reactions, an upstream mutagenesis primer (5′-TTCCAGATCTACGCTGTGCCTTTTCAGGG-3′ for W106F-cystatin C or 5′-TTCCAGATCTACGCTGTGCCTGGGCAGGG-3′ for W106G-cystatin C), designed to include the *Bgl*II site and span over the codon for cystatin C residue 106, was used for PCR with pHD313 as template. The same downstream primer (5′-CGACGTTGTAAAACGACGGC-3′), corresponding to a vector sequence downstream from the *Eco*RI site, was used in both reactions. The PCR amplifications were done in a Perkin-Elmer Cetus DNA Thermal Cycler with reagents from the AmpliTaq kit (Perkin-Elmer Cetus, Norwalk, CT), primers at 0.6 $\mu$M concentration, and 0.6 ng of uncleaved pHD313 DNA as template in a reaction volume of 100 $\mu$L. The PCR cycle (94 °C, 1 min−60 °C, 1 min−72 °C, 1 min) was repeated 30 times. The purified PCR products were cleaved with *Bgl*II/*Eco*RI (Life Technologies, Gaithersburg, MD), and ligated into *Bgl*II/*Eco*RI cut and dephosporylated pHD313. The expression plasmids mutated in this manner were introduced in *E. coli* strain MC1061, and the constructs were verified by DNA sequencing of plasmids isolated from bacterial subclones. To this end, a 569 bp segment including the entire coding sequence for cystatin C was amplified and sequenced as described earlier (Hall et al., 1995). The conditions for expression of the two cystatin C variants and the osmotic shock procedure used to collect the periplasmic fraction containing the recombinant proteins were as for expression of wild-type cystatin C (Lindahl et al., 1992a).

The two cystatin C variants and recombinant wild-type human cystatin C (Abrahamson et al., 1988) were purified by ion-exchange chromatography on Q-Sepharose (Pharmacia Biotech, Uppsala, Sweden), followed by gel chromatography on Sephadex G 75 (Pharmacia Biotech), as detailed earlier (Abrahamson et al., 1988; Lindahl et al., 1992a; Hall et al., 1995). Forms 1 and 2 of chicken cystatin were obtained from egg white (Anastasi et al., 1983; Lindahl et al., 1988).

An N-terminally truncated form of W106G-cystatin C with Gly-11 as the N-terminal residue (Gly$^{11}$-W106G-cystatin C) was obtained by digestion with neutrophil elastase (Abrahamson et al., 1991). The digestion product had the expected N-terminal sequence and showed only one band with an apparent molecular weight only slightly lower than that of the intact protein in SDS−PAGE under reducing conditions, indicating that no internal cleavages had occurred.

The purification and activation of papain (EC 3.4.22.2), actinidin (EC 3.4.22.14), and recombinant rat cathepsin B (EC 3.4.22.1), expressed in yeast, have been reported elsewhere (Lindahl et al., 1988; Björk & Ylinenjärvi, 1990; Björk et al., 1994). These studies showed that the enzyme preparations were essentially fully active in binding to wild-type cystatin C. Cathepsin L (EC 3.4.22.15) was from sheep liver (Mason, 1986). Human liver cathepsin H (EC 3.4.22.16) was purchased from Calbiochem (San Diego, CA); its specific activity was 0.2 unit/mg of protein with L-arginyl-2-naphthylamide as substrate at 40 °C, pH 6.8. The inactivated papain forms, *S*-(methylthio)papain, *S*-(carbamoylmethyl)papain, and *S*-(carboxymethyl)papain, were obtained as in previous work (Björk & Ylinenjärvi, 1989; Lindahl et al., 1992a).

*Spectroscopy.* Near-ultraviolet absorption difference spectra, near-ultraviolet circular dichroism spectra, and fluorescence emission spectra of W106G-cystatin C, papain, and the inhibitor−proteinase complex were measured as described previously (Lindahl et al., 1988, 1992a). Difference spectra were calculated from measured circular dichroism and fluorescence emission spectra as in Lindahl et al. (1988).

Titrations of active papain and *S*-(methylthio)papain with the two cystatin C variants for determination of binding stoichiometries and of *S*-(carbamoylmethyl)- and *S*-(carboxymethyl)papain with the variants for affinity measurements were monitored by the changes of fluorescence emission intensity induced by the binding (Lindahl et al., 1992a,b; Björk et al., 1995). The stoichiometry titrations were done at papain concentrations of 1 $\mu$M, whereas concentrations comparable to the dissociation constants were used in the affinity titrations. Stoichiometries and dissociation constants were obtained by nonlinear least-squares regression analysis of the titration curves (Björk et al., 1995).

*Inhibition Constants.* Inhibition constants for the interaction of wild-type cystatin C and the two cystatin C variants with actinidin or cathepsins B, H, and L were obtained from the equilibrium rates of cleavage of a fluorogenic substrate by the enzyme at increasing concentrations of the inhibitor, as detailed earlier (Björk et al., 1994, 1995; Pol et al., 1995). The substrate was carbobenzoxy-L-phenylalanyl-L-arginine 4-methylcoumaryl-7-amide (Peptide institute, Osaka, Japan) for actinidin and cathepsin L, carbobenzoxy-L-arginyl-L-arginine 4-methylcoumaryl-7-amide (Peptide Institute) for cathepsin B, and L-arginine 4-methylcoumaryl-7-amide (Bachem, Bubendorf, Switzerland) for cathepsin H. The substrate concentrations were 5−20 $\mu$M; substrate hydrolysis never exceeded 5%. The inhibitor concentration was varied from $\leq 0.5K_{i,app.}$ (the apparent inhibition constant) to $\geq 4K_{i,app.}$, except in the studies with cathepsin L. The lowest inhibitor concentration that could be analyzed with this enzyme was still well above $K_{i,app.}$, and only upper limits for the inhibition constants could therefore be obtained. The enzyme concentration was at least 10-fold lower than that of the inhibitor in all experiments. Product formation was monitored continuously in the fluorometer, in most cases until equilibrium was reached. For the slowest reactions, however, the equilibrium rates of substrate hydrolysis were obtained by nonlinear regression analysis of the progress curves (Björk et al., 1994). The $K_m$ values used for correction of the apparent inhibition constants for substrate competition have been reported elsewhere (Schwartz & Barrett, 1980; Mason, 1986; Lindahl et al., 1992b; Björk et al., 1994, 1995).

The inhibition constants for the interaction of Gly$^{11}$-W106G-cystatin C with papain, cathepsin B, and cathepsin L were determined in essentially the same manner. However, carbobenzoxy-L-arginine 4-methylcoumaryl-7-amide (Sigma, St. Louis, MO) was used as substrate for papain and cathepsin B, and carbobenzoxy-L-arginyl-L-arginine 4-

[1] Abbreviations: DTT, dithiothreitol; E-64, 1-[[*N*-(L-3-*trans*-carboxyoxirane-2-carbonyl)-L-leucyl]amino]-4-guanidinobutane; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; W106F- and W106G-cystatin C, cystatin C variants in which Trp-106 is replaced with Phe and Gly, respectively; Gly$^{11}$-W106G-cystatin C, a truncated form of W106G-cystatin C having Gly-11 as its N-terminal amino acid.

methylcoumaryl-7-amide was used for cathepsin L. These less efficient substrates allowed the molar ratio of inhibitor to enzyme to be fixed at 10:1 throughout the concentration range covered, thus minimizing the effect of a possible small contamination of intact W106G-cystatin C in the preparation of the truncated form. The low affinity of $Gly^{11}$-W106G-cystatin C for cathepsin B precluded analyses at higher inhibitor concentrations than about $K_{i,app}$. The $K_m$ values used for correction of the apparent inhibition constants are from previous work (Roberts et al., 1986; Dalet-Fumeron et al., 1991; Hasnain et al., 1992).

*Association Kinetics.* The kinetics of binding of the two cystatin C variants to active papain or *S*-(methylthio)papain were studied by stopped-flow fluorescence measurements under pseudo-first-order conditions in an Applied Photophysics (Leatherhead, U.K.) SX-17MV instrument, essentially as in Björk et al. (1989). The inhibitor concentration was varied between 0.5 and 6 $\mu$M, and the enzyme concentration was maintained 10-fold lower.

The rate of binding of wild-type cystatin C and the cystatin C variants to cathepsins B and L was analyzed under pseudo-first-order conditions by continuous measurements of the loss of enzyme activity in the presence of a fluorogenic substrate, essentially as in earlier work (Björk et al., 1994, 1995; Pol et al., 1995). The substrates, their concentrations, and maximal substrate hydrolysis were the same as in the measurements of inhibition constants. The inhibitor concentration was maintained at least 10-fold higher than that of the enzyme and was varied from 0.5 to 2.5 $\mu$M for cathepsin B and from 0.2 to 2 nM for cathepsin L. The reactions with cathepsin L were monitored by conventional fluorometry, whereas the reactions with cathepsin B had to be analyzed by stopped-flow fluorometry, due to higher rates caused by a higher dissociation rate constant. Pseudo-first-order rate constants were evaluated by nonlinear regression analysis of the progress curves (Björk et al., 1994). The apparent second-order rate constants were corrected for substrate competition as described for the determination of inhibition constants.

*Dissociation Kinetics.* The rate of dissociation of complexes between the two cystatin C variants and papain was evaluated by trapping the enzyme dissociated from the complex (10 $\mu$M) by an excess (100−200 $\mu$M) of chicken cystatin (form 2), as in previous studies with wild-type cystatin C (Lindahl et al., 1992a). The dissociation rate constant was obtained from the initial rate of appearance of the complex between the displacing cystatin 2 and the liberated papain, monitored by chromatography on a Mono-Q (Pharmacia Biotech) ion-exchange column.

The kinetics of dissociation of complexes between the cystatin C variants and *S*-(methylthio)papain were analyzed by similarly trapping the inactivated enzyme liberated from 1 $\mu$M complex by 2−12 $\mu$M chicken cystatin (form 1). The rate of the reaction was monitored by the decrease of tryptophan fluorescence caused by the tight binding of the displacing chicken cystatin to the liberated *S*-(methylthio)-papain, as in previous work (Pol et al., 1995).

*Experimental Conditions and Protein Concentrations.* All analyses were done at 25.0 ± 0.2 °C. Reactions with active papain, inactivated papains, and actinidin were analyzed in 0.05 M Tris-HCl, pH 7.4, 0.1 M NaCl, and 100 $\mu$M EDTA. The buffers in studies of reactions with other enzymes were as follows: for cathepsin B, 0.05 M Mes, pH 6.0, 0.1 M NaCl, 100 $\mu$M EDTA, 0.5 mM DTT, and 0.1% (w/v) poly(ethylene glycol); for cathepsin H, 0.1 M sodium phosphate, pH 6.0, 1 mM EDTA, and 1 mM DTT; for cathepsin L, 0.1 M sodium acetate, pH 5.5, 1 mM EDTA, 1 mM DTT, and 0.005% (w/v) Brij 35.

Protein concentrations were obtained by absorption measurements at 280 nm. A molar absorption coefficient of $1.11 \times 10^4$ $M^{-1}$ $cm^{-1}$ was determined previously for wild-type cystatin C (Lindahl et al., 1992a). A molar absorption coefficient of $5.4 \times 10^3$ $M^{-1}\cdot cm^{-1}$ was calculated for W106F-, W106G-, and $Gly^{11}$-W106G-cystatin C from this value by subtracting the molar absorption coefficient of one tryptophan residue (Gill & von Hippel, 1989). Molar absorption coefficients of $5.59 \times 10^4$, $4.98 \times 10^4$, and $6.34 \times 10^4$ $M^{-1}\cdot cm^{-1}$ were used for papain, actinidin, and cathepsin B (Lindahl et al., 1988; Björk & Ylinenjärvi, 1990; Björk et al., 1994). The concentration and relative molecular mass of human cathepsin H were provided by the manufacturer. The molar concentration used for sheep cathepsin L was determined by active site-titration with E-64 (Mason, 1986).

## RESULTS

The importance of the second hairpin loop of cystatin C for target enzyme binding was elucidated by studies of two inhibitor variants with the conserved Trp-106 residue replaced with Phe or Gly, produced by site-directed mutagenesis and *E. coli* expression. Sequencing of the entire coding region of the expression vectors verified that the correct mutations and no other sequence alterations had been introduced. The purified proteins showed only one band in SDS−PAGE under reducing conditions (not shown).

*Stoichiometry of Binding of Cystatin C Variants to Papain and Spectroscopic Characterization of the Interaction.* Titrations of active papain and its inactive derivative, *S*-(methylthio)papain, with the two cystatin C variants for determination of binding stoichiometries were monitored by the changes in intrinsic fluorescence accompanying the interaction (see below). Stoichiometries of 1.08 ± 0.03 (SE; $n = 4$) and 1.04 ± 0.04 ($n = 4$) were obtained for W106F- and W106G-cystatin C, respectively, showing that both variants were essentially fully active in binding proteinases.

The changes of near-ultraviolet absorption, circular dichroism, and fluorescence induced by the binding of wild type and W106G-cystatin C to papain were compared (Figure 1). Both the ultraviolet absorption difference spectrum and the near-ultraviolet circular dichroism difference spectrum for the W106G-cystatin C−papain interaction were markedly different from the corresponding spectra for wild-type cystatin C. The fluorescence emission difference spectra for the variant and wild-type inhibitors also differed, although to a lesser extent. The fluorescence emission spectra of both W106F- and W106G-cystatin C showed maxima around 310 nm, confirming the absence of tryptophan residues.

*Equilibrium and Kinetic Characterization of Interactions of Cystatin C Variants with Cysteine Proteinases.* Dissociation equilibrium constants, as well as association and dissociation rate constants, were measured for the binding of W106F-, W106G-, and $Gly^{11}$-W106G-cystatin C to several cysteine proteinases (Table 1). Most dissociation equilibrium constants were determined as inhibition constants. However, the dissociation constants for the tight binding to papain were

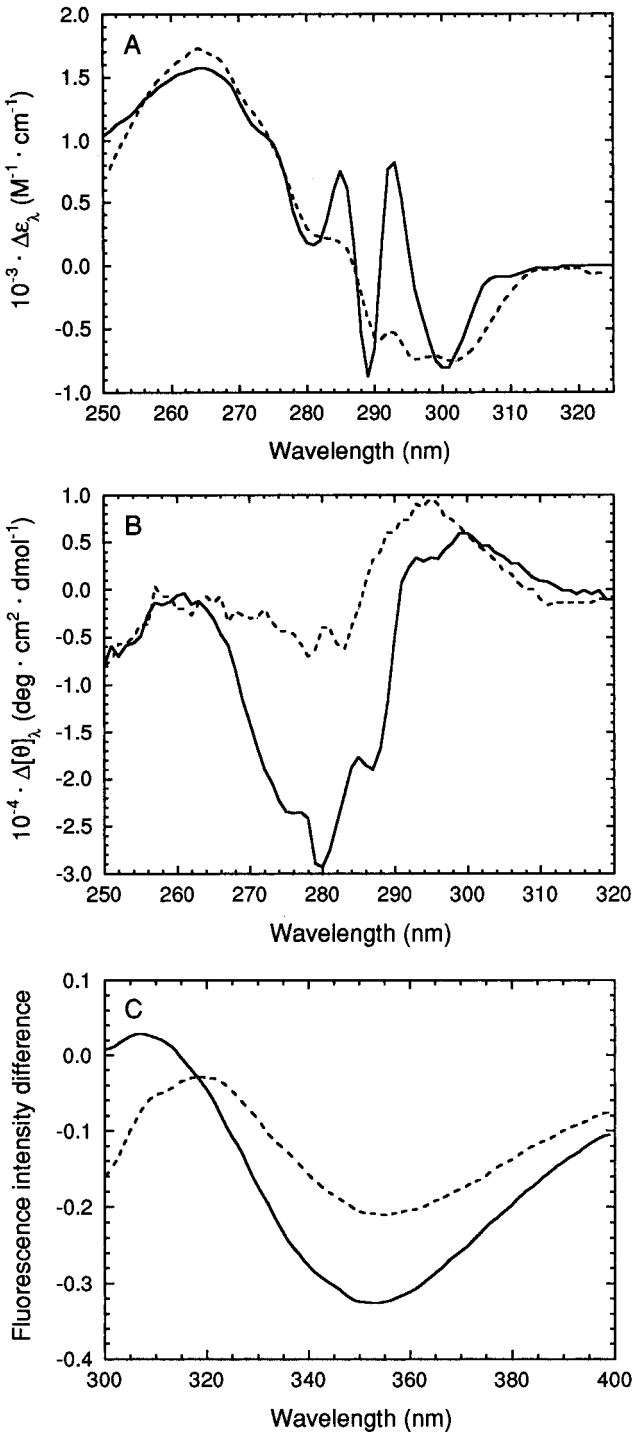


FIGURE 1: Spectroscopic changes on the interaction of wild-type or W106G-cystatin C with papain. (A) Near-ultraviolet absorption difference spectra. (B) Near-ultraviolet circular dichroism difference spectra. (C) Fluorescence emission difference spectra. (—) Wild-type cystatin C; (---) W106G-cystatin C. The papain concentrations were 15 $\mu$M in (A) and (B) and 1 $\mu$M in (C). The molar ratio of cystatin to papain was 1.5, giving >99% saturation of the enzyme. The unit on the ordinate in (B) is the difference in molar ellipticity. The data for wild-type cystatin C, shown for comparison, are from Lindahl et al. (1992a).

not measured directly but were calculated from separately determined association and dissociation rate constants. Moreover, only upper limits for the inhibition constants for cathepsin L could be estimated, owing to the tight interaction. Temporary inhibition, as reported for variants of chicken cystatin with the second hairpin loop deleted (Machleidt et al., 1995), was not apparent for any of the enzymes. Association rate constants for reactions with papain were monitored by tryptophan fluorescence in a stopped-flow instrument. The corresponding rate constants for reactions with cathepsin B and cathepsin L were obtained from the loss of enzyme activity in the presence of a fluorogenic substrate, analyzed by stopped-flow fluorometry for cathepsin B and by conventional fluorometry for cathepsin L. Rate constants for reactions with actinidin or cathepsin H could not be obtained by either method, due to low affinities of the interactions, the lack of sufficiently good substrates for the enzymes, and, in the case of cathepsin H, also to the small amounts of enzyme available. Most dissociation rate constants were calculated from measured dissociation equilibrium and association rate constants. Dissociation rate constants for the tight interactions with papain, however, were measured by trapping the dissociated enzyme with an excess of chicken cystatin and monitoring the appearance of trapped enzyme. The amounts of cathepsin L at hand were insufficient to allow similar measurements for interactions with this enzyme.

*Interaction of Cystatin C Variants with Inactivated Papains.* Dissociation equilibrium constants for the binding of the cystatin C variants to *S*-(methylthio)papain were calculated from association constants, measured by stopped-flow fluorescence, and dissociation constants, measured by a displacement technique (Table 2). Dissociation equilibrium constants for the weaker interactions of the variants with *S*-(carbamoylmethyl)- and *S*-(carboxymethyl)papain could be measured directly by titrations monitored by the decrease of tryptophan fluorescence induced by the binding (Table 2). The dissociation constant for the binding of wild-type cystatin C to *S*-(carboxymethyl)papain, needed for comparison (Table 2), is not available in the literature and was found to be of a magnitude such that it could not be determined by any of the methods used for other interactions.

## DISCUSSION

The markedly altered spectroscopic changes on papain binding that resulted from substitution of the single Trp residue of cystatin C, Trp-106, by Gly show that this residue is involved in the interaction with the proteinase, in agreement with previous evidence (Lindahl et al., 1988; Bode et al., 1988; Nycander & Björk, 1990). Only a small change of near-ultraviolet circular dichroism was induced by the Gly variant, showing that the circular dichroism changes accompanying the binding of the wild-type inhibitor arise predominantly from Trp-106 (Lindahl et al., 1988, 1992a). The nature of the spectroscopic changes caused by the binding of the Gly variant indicates that also these changes arise predominantly from Trp residues. These tryptophans must be located in papain, likely candidates being Trp-177 and Trp-69 near the active site of the enzyme, consistent with previous suggestions (Lindahl et al., 1988; Bode et al., 1988). The spectroscopic changes induced by binding of the Gly variant to papain were highly similar to those induced by the related cystatin, cystatin A (Pol et al., 1995), which also lacks tryptophan. This observation indicates that most of the binding surfaces of cystatins A and C interact with papain in a similar overall manner.

Substitution of Trp-106 of cystatin C by Phe was found to lead to a marked loss of the affinity of the inhibitor for

Table 1: Dissociation Equilibrium Constants ($K_d$), Association Rate Constants ($k_{ass}$), and Dissociation Rate Constants ($k_{diss}$) for the Interactions of Wild-Type, Trp-106 Variants, and an N-Terminally Truncated Trp-106 Variant of Cystatin C with Papain, Actinidin, and Cathepsins B, H, and L[a]

| proteinase | cystatin C form | $K_d$ (M) | $k_{ass}$ ($M^{-1}\cdot s^{-1}$) | $k_{diss}$ ($s^{-1}$) |
|---|---|---|---|---|
| papain | wild-type | $1.1 \times 10^{-14}$ [b] | $1.1 \times 10^{7}$ [b] | $1.3 \times 10^{-7}$ [b] |
| | W106F | $7.5 \times 10^{-13}$ [c] | $(3.6 \pm 0.4) \times 10^{6}$ ($n = 8$) | $(2.7 \pm 0.3) \times 10^{-6}$ ($n = 3$) |
| | W106G | $9.6 \times 10^{-12}$ [c] | $(2.8 \pm 0.1) \times 10^{6}$ ($n = 9$) | $(2.7 \pm 0.1) \times 10^{-5}$ ($n = 3$) |
| | $Gly^{11}$-W106G | $(9.0 \pm 0.8) \times 10^{-9}$ ($n = 6$) | ND | ND |
| actinidin | wild-type | $1.9 \times 10^{-8}$ [b] | $2.4 \times 10^{6}$ [b] | $4.6 \times 10^{-2}$ [b] |
| | W106F | $(2.5 \pm 0.2) \times 10^{-6}$ ($n = 9$) | ND | ND |
| | W106G | $(6.6 \pm 0.6) \times 10^{-6}$ ($n = 11$) | ND | ND |
| cathepsin B | wild-type | $3.2 \times 10^{-10}$ [b] | $1.1 \times 10^{6}$ [b] | $3.5 \times 10^{-4}$ [b] |
| | W106F | $(3.9 \pm 0.1) \times 10^{-8}$ ($n = 9$) | $(9.7 \pm 0.2) \times 10^{5}$ ($n = 5$) | $3.8 \times 10^{-2}$ [c] |
| | W106G | $(9.7 \pm 0.3) \times 10^{-8}$ ($n = 9$) | $(1.2 \pm 0.1) \times 10^{6}$ ($n = 5$) | 0.12[c] |
| | $Gly^{11}$-W106G | $(1.1 \pm 0.2) \times 10^{-5}$ ($n = 5$) | ND | ND |
| cathepsin H | wild-type | $(4.2 \pm 0.3) \times 10^{-10}$ ($n = 14$) | ND | ND |
| | W106F | $(1.3 \pm 0.1) \times 10^{-8}$ ($n = 12$) | ND | ND |
| | W106G | $(2.6 \pm 0.1) \times 10^{-7}$ ($n = 16$) | ND | ND |
| cathepsin L | wild-type | $<5 \times 10^{-12}$ | $(2.3 \pm 0.2) \times 10^{7}$ ($n = 10$) | $<1 \times 10^{-4}$ [c] |
| | W106F | $<5 \times 10^{-12}$ | $(1.2 \pm 0.1) \times 10^{7}$ ($n = 10$) | $<6 \times 10^{-5}$ [c] |
| | W106G | $\leq 5 \times 10^{-12}$ | $(7.8 \pm 0.6) \times 10^{6}$ ($n = 10$) | $\leq 4 \times 10^{-5}$ [c] |
| | $Gly^{11}$-W106G | $(7.8 \pm 0.5) \times 10^{-9}$ ($n = 7$) | ND | ND |

[a] The values for wild-type cystatin C, some of which are from previous work (Lindahl et al., 1992a; Björk et al., 1994), are shown for comparison. Values measured in this work are given with their standard errors and the number of measurements in parentheses. ND, not determined. [b] From previous work. [c] Calculated value.

Table 2: Dissociation Equilibrium Constants (in M) for the Interactions of Wild-Type and Trp-106 Variants of Cystatin C with Inactivated Forms of Papain[a]

| | inactivating group | | | |
|---|---|---|---|---|
| cystatin form | none | *S*-(methylthio) | *S*-(carbamoylmethyl) | *S*-(carboxymethyl) |
| wild-type | $1.1 \times 10^{-14}$ [b] | $2.2 \times 10^{-13}$ [b] | $8.3 \times 10^{-11}$ [b] | ND |
| W106F | $7.5 \times 10^{-13}$ [c] | $5.5 \times 10^{-11}$ [d] | $(1.0 \pm 0.2) \times 10^{-8}$ ($n = 3$) | $3.6 \times 10^{-7}$ ($n = 1$) |
| W106G | $9.6 \times 10^{-12}$ [c] | $8.7 \times 10^{-10}$ [e] | $(8.5 \pm 1.1) \times 10^{-8}$ ($n = 4$) | $1.6 \times 10^{-6}$ ($n = 1$) |

[a] The values for wild-type cystatin C, which are from previous work (Lindahl et al., 1992a; Björk et al., 1995), and the values for active papain (from Table 1) are shown for comparison. Values measured in this work are given with their standard errors, if more than one measurement was made, and with the number of measurements in parentheses. ND, not determined. [b] From previous work. [c] From Table 1. [d] Calculated from an association rate constant of $(3.6 \pm 0.2) \times 10^{6}$ $M^{-1}\cdot s^{-1}$ ($n = 6$) and a dissociation rate constant of $(2.0 \pm 0.3) \times 10^{-4}$ $s^{-1}$ ($n = 5$). [e] Calculated from an association rate constant of $(3.1 \pm 0.3) \times 10^{6}$ $M^{-1}\cdot s^{-1}$ ($n = 7$) and a dissociation rate constant of $(2.7 \pm 0.1) \times 10^{-3}$ $s^{-1}$ ($n = 5$).

cysteine proteinases, ~30-fold for cathepsin H, ~70-fold for papain, and ~120-fold for actinidin and cathepsin B. Substitution by Gly lead to a further, although substantially smaller, decrease in affinity, ~2.5-fold for actinidin and cathepsin B, ~10-fold for papain, and ~20-fold for cathepsin H. The Gly mutant had a 300−900-fold lower affinity for the enzymes than the wild-type inhibitor, in reasonable agreement with earlier work (Hall et al., 1995). Unitary free energy changes (i.e., the free energy changes after elimination of the decrease in entropy resulting from two molecules forming one; Gurney, 1953; Karush, 1962) were calculated for the different interactions from the measured dissociation equilibrum constants. These values indicate that the side chain of Trp-106 contributes a similar free energy, −14 to −17 kJ·mol$^{-1}$, to the binding to all four cysteine proteinases for which data could be obtained. This free energy corresponds to about 20−30% of the total energy of binding of cystatin C to the proteinases, which ranges from −90 kJ·mol$^{-1}$ for papain to −54 kJ·mol$^{-1}$ for actinidin. The similar contribution from Trp-106 shows that these latter differences in binding energies originate from other interactions than those involving this residue. Comparison with previous data for intact and N-terminally truncated cystatin C (Björk et al., 1994) indicates that Trp-106 provides about half the free energy of the N-terminal region in the binding of papain. In contrast, Trp-106 and the N-terminal region contribute comparable binding energies to the interactions with actinidin or cathepsin B, in agreement with previous results (Hall et al., 1995).

The calculations of unitary free energy changes also show that, in contrast to the comparatively constant binding energy arising from Trp-106, a Phe residue in this position contributes a varying, but considerably smaller, energy of binding to the four proteinases studied, about −2.5 kJ·mol$^{-1}$ for actinidin and cathepsin B and about −7 kJ·mol$^{-1}$ for papain and cathepsin H. Depending on the enzyme, these free energies correspond to only 20−45% of the binding energy of a Trp residue, the largest contribution being apparent for cathepsin H. Hydrophobic interactions involving a phenyl group thus cannot substitute for those involving an indole ring in position 106, although they can provide part of the binding energy of the latter.

Characterization of the binding of the cystatin C variants to inactivated papains revealed a similar pattern. By analogous free energy calculations, Trp-106 was found to contribute −17 to −20 kJ·mol$^{-1}$ to the interaction with *S*-(methylthio)- and *S*-(carbamoylmethyl)papain, i.e., about 25% of the total energy of binding of cystatin C to these inactivated enzymes. In contrast, a Phe residue contributes only −5 to −7 kJ·mol$^{-1}$, corresponding to about 30% of the binding energy of tryptophan. An inctivating group on the active-site cysteine of papain, although reducing the overall affinity for cystatins (Björk & Ylinenjärvi, 1989), apparently does not appreciably alter the interactions between

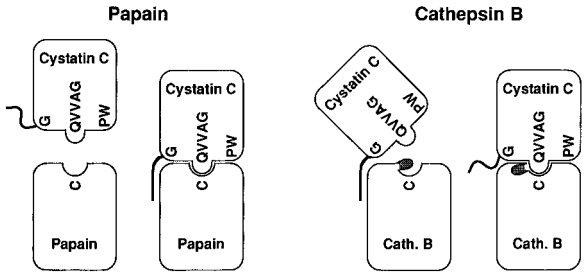


FIGURE 2: Schematic model of the proposed different roles of the N-terminal region and the first and second hairpin loops of cystatin C in the inactivation of papain and cathepsin B. In the reaction with papain (left), the flexible N-terminal region (thick line), extending from Gly-11 (G), interacts with the enzyme concurrent with, or at a rate indistinguishable from that of, the first hairpin loop (QVVAG) and the second hairpin loop (PW). Moreover, all three regions keep the inhibitor attached to the proteinase in the complex. In the reaction with cathepsin B (right), the N-terminal region is the first to interact with the enzyme and orients the inhibitor so that the "occluding loop" (shaded) is displaced and the first and second hairpin loops can bind. However, the N-terminal region does not contribute appreciable binding energy once the complex has been formed. The active-site cysteine of the proteinase is denoted by C.

the second hairpin loop of the inhibitor and the enzyme.

The decreased affinities between the cystatin C Trp-106 variants and all active or inactivated enzymes investigated were due almost exclusively to increased dissociation rate constants, the association rate constants being minimally affected. This behavior is in contrast to that observed for substitutions in the N-terminal regions of cystatin C or chicken cystatin and for truncation of these regions by proteolysis (Abrahamson et al., 1991; Björk et al., 1994, 1995). Such changes reduced the affinity of the inhibitors for proteinases by increasing the dissociation rate constants for most enzymes, but by decreasing the association rate constants for cathepsin B. These and other observations indicate that the N-terminal region and the second hairpin loop of cystatin C both contribute to the affinity for most enzymes primarily by keeping enzyme and inhibitor attached in the complex, the rate of association being largely diffusion-controlled (Björk et al., 1989, 1994, 1995; Björk & Ylinen-järvi, 1990; Lindahl et al., 1992a; Figure 2, left). In contrast, the two regions appear to have different functions in the interaction of cystatin C with cathepsin B (Björk et al., 1994, 1995). The main role of the N-terminal region thus presumably is to bind to this proteinase before the remainder of the binding surface of the inhibitor and to properly orient the latter with respect to the enzyme (Björk et al., 1994; Figure 2, right). In this way, the rate of the reaction is increased by the inhibitor being able to rotate and tilt in such a manner that it can efficiently displace the "occluding loop", which partially restricts the access to the active site of cathepsin B (Musil et al., 1991). Once the complex with cathepsin B has been formed, however, the N-terminal region does not contribute appreciably to its stability (Björk et al., 1994). The second hairpin loop, on the other hand, appears to function primarily by keeping cystatin C anchored to cathepsin B in the complex in a similar manner as in the case of interactions with other proteinases. The first hairpin loop, i.e., the "QVVAG-region", presumably stabilizes proteinase complexes analogous to the second hairpin loop (Auerswald et al., 1995), although sufficient kinetic data to fully verify this conclusion are lacking.

As expected (Abrahamson et al., 1987, 1991; Lindahl et al., 1992b; Björk et al., 1994), removal of the N-terminal region of the Trp-106→Gly variant by proteolytic cleavage resulted in a substantial decrease in the affinity of this variant for the three enzymes tested, papain, cathepsin B, and cathepsin L. The unitary free energy of binding of the truncated Gly variant to both papain and cathepsin B was about 60% of that of the wild-type; the decrease for cathepsin L could not be quantified, due to lack of a value for the affinity of the wild-type inhibitor for this enzyme. Proteolytic truncation of the N-terminal region of cystatin C and chicken cystatin has been shown previously essentially to eliminate all interactions of this region with proteinases, without appreciably affecting the binding of the other parts of the binding surface (Lindahl et al., 1992b; Björk et al., 1994). Similarly, substitution of Trp-106 by Gly may eliminate most or all interactions of the second hairpin loop, while minimally interfering with the binding of the remainder of the inhibitor. Such an assumption leads to the conclusion that the first hairpin loop (the "QVVAG-region"), which would then be the only region of the N-terminally truncated Gly variant interacting with the proteinase, is responsible for about 60% of the energy of binding of cystatin C to both papain and cathepsin B. This value agrees well with the contribution by the first hairpin loop to cathepsin B binding of about 60% estimated from the binding energy of the wild-type inhibitor and the independently evaluated binding energies of the N-terminal region (Björk et al., 1994) and Trp-106. However, the value for papain is higher than that of about 40% estimated for this enzyme in the same manner. Certain interactions with papain involving the second hairpin loop may thus remain on Trp-106 substitution, or the first hairpin loop may bind somewhat differently to papain in the absence of the other two binding regions.

## ACKNOWLEDGMENT

We are grateful to Drs. John S. Mort, Shriners Hospital, Montreal, Canada, and Robert W. Mason, Alfred I. duPont Institute, Wilmington, DE, for gifts of recombinant rat cathepsin B and sheep cathepsin L, respectively.

## REFERENCES

Abrahamson, M., Ritonja, A., Brown, M. A., Grubb, A., Machleidt, W., & Barrett, A.J. (1987) *J. Biol. Chem. 262*, 9688−9694.
Abrahamson, M., Dalbøge, H., Olafsson, I., Carlsen, S., & Grubb, A. (1988) *FEBS Lett. 236*, 14−18.
Abrahamson, M., Mason, R. W., Hansson, H., Buttle, D. J., Grubb, A., & Ohlsson, K. (1991) *Biochem. J. 273*, 621−626.
Anastasi, A., Brown, M. A., Kembhavi, A. A., Nicklin, M. J., Sayers, C. A., Sunter, D. C., & Barrett, A. J. (1983) *Biochem. J. 211*, 129−138.
Auerswald, E. A., Genenger, G., Assfalg-Machleidt, I., Machleidt, W., Engh, R. A., & Fritz, H. (1992) *Eur. J. Biochem. 209*, 837−845.
Auerswald, E. A., Nägler, D. K., Schulze, A. J., Engh, R. A., Genenger, G., Machleidt, W., & Fritz, H. (1994) *Eur. J. Biochem. 224*, 407−415.
Auerswald, E. A., Nägler, D. K., Assfalg-Machleidt, I., Stubbs, M. T., Machleidt, W., & Fritz, H. (1995) *FEBS Lett. 361*, 179−184.
Björk, I., & Ylinenjärvi, K. (1989) *Biochem. J. 260*, 61−68.
Björk, I., & Ylinenjärvi, K. (1990) *Biochemistry 29*, 1770−1776.
Björk, I., Alriksson, E., & Ylinenjärvi, K. (1989) *Biochemistry 28*, 1568−1573.
Björk, I., Pol, E., Raub-Segall, E., Abrahamson, M., Rowan, A. D., & Mort, J. S. (1994) *Biochem. J. 299*, 219−225.

Björk, I., Brieditis, I., & Abrahamson, M. (1995) *Biochem. J. 306*, 513−518.
Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., & Turk, V. (1988) *EMBO J. 7*, 2593−2599.
Dalet-Fumeron, V., Guinec, N., & Pagano, M. (1991) *J. Chromatogr. 568*, 55−68.
Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
Gurney, R. W. (1953) in *Ionic Processes in Solution*, pp 90−105, McGraw-Hill, New York.
Hall, A., Dalbøge, H., Grubb, A., & Abrahamson, M. (1993) *Biochem. J. 291*, 123−129.
Hall, A., Håkansson, K., Mason, R. W., Grubb, A., & Abrahamson, M. (1995) *J. Biol. Chem. 270*, 5115−5121.
Hasnain, S., Hirama, T., Tam, A., & Mort, J. S. (1992) *J. Biol. Chem. 267*, 4713−4721.
Karush, F. (1962) *Adv. Immunol. 2*, 1−40.
Lindahl, P., Alriksson, E., Jörnvall, H., & Björk, I. (1988) *Biochemistry 27*, 5074−5082.
Lindahl, P., Abrahamson, M., & Björk, I. (1992a) *Biochem. J. 281*, 49−55.
Lindahl, P., Nycander, M., Ylinenjärvi, K., Pol, E., & Björk, I. (1992b) *Biochem. J. 286*, 165−171.
Machleidt, W., Thiele, U., Laber, B., Assfalg-Machleidt, I., Esterl, A., Wiegand, G., Kos, J., Turk, V., & Bode, W. (1989) *FEBS Lett. 243*, 234−238.
Machleidt, W., Nägler, D. K., Assfalg-Machleidt, I., Stubbs, M. T., Fritz, H., & Auerswald, E. A. (1995) *FEBS Lett. 361*, 185−190.
Mason, R. W. (1986) *Biochem. J. 240*, 285−288.
Musil, D., Zučič, D., Turk, D., Engh, R. A., Mayr, I., Huber, R., Popovič, T., Turk, V., Towatari, T., & Katunuma, N. (1991) *EMBO J. 10*, 2321−2330.
Nycander, M., & Björk, I. (1990) *Biochem. J. 271*, 281−284.
Pol, E., Olsson, S. L., Estrada, S., Prasthofer, T. W., & Björk, I. (1995) *Biochem. J. 311*, 275−282.
Roberts, D. D., Lewis, S. D., Ballou, D. P., Olson, S. T., & Shafer, J. A. (1986) *Biochemistry 25*, 5595−5601.
Schwartz, W. N., & Barrett, A. J. (1980) *Biochem. J. 191*, 487−497.
Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarčič, B., & Turk, V. (1990) *EMBO J. 9*, 1939−1947.

BI960420U