ACTA UNIVERSITATIS AGRICULTURAE SUECIAE

VETERINARIA 39

Cystatin A, a mammalian cysteine proteinase inhibitor

Mechanism of inhibition of target proteinases by recombinant cystatin A variants

Sergio Estrada

SWEDISH UNIVERSITY OF AGRICULTURAL SCIENCES

Cystatin A, a mammalian cysteine proteinase inhibitor. Mechanism of inhibition of target proteinases by recombinant cystatin A variants

Sergio Estrada

Akademisk avhandling som för vinnande av filosofie doktorsexamen kommer att offentligen försvaras i Ettans föreläsningssal, Klinikcentrum, SLU, Uppsala, fredagen den 25 september 1998, kl. 9.15.

Abstract

Cystatin A, a mammalian cysteine proteinase inhibitor, was expressed in a bacterial system. The purified protein was fully active. It inhibited the proteinases,papain and cathepsin L, strongly, with K_i values of 0.2–20 pM, whereas the affinities for actinidin and cathepsins B, C and H were in the range of 1–40 nM. The binding to papain and cathepsin L was rapid, with k_{ass} ~3–5 × 10^6 $M^{-1} \cdot s^{-1}$, whereas the inhibition of cathepsin B was ~100-fold slower. The binding to papain and consistent with a one-step binding mechanism.

Gly-4 mutants of cystatin A had lower affinities for papain and cathepsins B and L than the wild-type inhibitor. In general, the K_i values increased with the size of the side chain of the mutant. Even the smallest mutation, Gly-4 to Ala, had substantial effects, and the charged Gly-4 to Glu or Arg variants had affinities for the enzymes that were more than five orders of magnitude lower than those of the wild-type. The rate of binding of the mutants to papain and cathepsin L was unaffected, but was lower than that of the wild-type for cathepsin B, presumably reflecting structural differences in the enzymes.

Sequential truncations of the N-terminal region of cystatin A showed that Ile-2 and, to an even greater extent, Pro-3 were the residues of the this region that contributed to the binding of the inhibitor to papain and cathepsins B and L. In contrast with the primarily anchoring role of the N-terminal region of cystatin A to papain and cathepsin L, this region was needed also for mantaining the rate of association of the inhibitor with cathepsin B.

A fluorescent probe was linked to a recombinant cystatin A variant via an extra N-terminal cysteine. The labelled inhibitor had eight-fold higher affinity for papain than the wild-type cystatin, due to an increased association rate constant. The binding of the N-terminal label followed saturation kinetics, indicating that the N-terminal region of the labelled cystatin binds to the enzyme in the second step of a two-step reaction mechanism.

Keywords: cysteine proteinase, cysteine proteinase inhibitor, papain, cathepsins, cystatins, enzyme kinetics, inhibition, recombinant protein, stefins, two-step reaction mechanism

Distribution:
Swedish University of Agricultural Sciences
Department of Veterinary Medical Chemistry
S-751 23 UPPSALA, Sweden

Uppsala 1998
ISSN 1401-6257
ISBN 91576 5448 4

Cystatin A, a mammalian cysteine proteinase inhibitor

Mechanism of inhibition of target proteinases by recombinant cystatin A variants

Sergio Estrada
Department of Veterinary Medical Chemistry
Uppsala

Doctoral thesis
Swedish University of Agricultural Sciences
Uppsala 1998

Acta Universitatis Agriculturae Sueciae
Veterinaria 39

ISSN 1401-6257
ISBN 91-576-5448-4

Tryck: SLU Service/Repro, Uppsala 1998

Abstract

Estrada, S., 1998. *Cystatin A, a mammalian cysteine proteinase inhibitor. Mechanism of inhibition of target proteinases by recombinant cystatin A variants.* Doctor's dissertation. ISSN 1401-6257, ISBN 91576 5448 4.

Cystatin A, a mammalian cysteine proteinase inhibitor, was expressed in a bacterial system. The purified protein was fully active. It inhibited the proteinases,papain and cathepsin L, strongly, with K_i values of 0.2–20 pM, whereas the affinities for actinidin and cathepsins B, C and H were in the range of 1–40 nM. The binding to papain and cathepsin L was rapid, with k_{ass} ~3–5 × 10^6 $M^{-1} \cdot s^{-1}$, whereas the inhibition of cathepsin B was ~100-fold slower. The binding to papain and consistent with a one-step binding mechanism.

Gly-4 mutants of cystatin A had lower affinities for papain and cathepsins B and L than the wild-type inhibitor. In general, the K_i values increased with the size of the side chain of the mutant. Even the smallest mutation, Gly-4 to Ala, had substantial effects, and the charged Gly-4 to Glu or Arg variants had affinities for the enzymes that were more than five orders of magnitude lower than those of the wild-type. The rate of binding of the mutants to papain and cathepsin L was unaffected, but was lower than that of the wild-type for cathepsin B, presumably reflecting structural differences in the enzymes.

Sequential truncations of the N-terminal region of cystatin A showed that Ile-2 and, to an even greater extent, Pro-3 were the residues of the this region that contributed to the binding of the inhibitor to papain and cathepsins B and L. In contrast with the primarily anchoring role of the N-terminal region of cystatin A to papain and cathepsin L, this region was needed also for mantaining the rate of association of the inhibitor with cathepsin B.

A fluorescent probe was linked to a recombinant cystatin A variant via an extra N-terminal cysteine. The labelled inhibitor had eight-fold higher affinity for papain than the wild-type cystatin, due to an increased association rate constant. The binding of the N-terminal label followed saturation kinetics, indicating that the N-terminal region of the labelled cystatin binds to the enzyme in the second step of a two-step reaction mechanism.

Keywords: cysteine proteinase, cysteine proteinase inhibitor, papain, cathepsins, cystatins, enzyme kinetics, inhibition, recombinant protein, stefins, two-step reaction mechanism

Authors address: Sergio Estrada, Department of Veterinary Medical Chemistry, SLU, Box 575, S-751 23 UPPSALA, Sweden.

Contents

Appendix

Papers on which this thesis is based

Papers I–IV

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals.

I. Pol, E., Olsson, S-L., Estrada, S., Prasthofer, T. and Björk, I. 1995. Characterization by spectroscopic, kinetic and equilibrium methods of the interaction between recombinant human cystatin A (stefin A) and cysteine proteinases. *Biochem. J. 311*, 275-282.

II. Estrada, S., Nycander, M., Hill, N., Craven, J., Waltho, J. and Björk, I. 1998. The role of human cystatin A (stefin A) in the binding of target proteinases. Characterization by kinetic and equilibrium methods of the interaction of cystatin A Gly-4 mutants with papain, cathepsin B and cathepsin L. *Biochemistry 37*, 7551-7560.

III. Estrada, S., Pavlova, A. and Björk, I. 1998. The importance of the N-terminal region of cystatin A for proteinase binding. Studies of the interaction between N-terminally truncated variants of cystatin A and papain, cathepsin B and cathepsin L. *Manuscript.*

IV. Estrada,, S. and Björk, I. 1998. Interaction of the N-terminal region of cystatin A with papain subsequent to binding of the haipin loops of the inhibitor. Rapid kinetic studies with cystatin A fluorescently labelled at tha N-terminus. *Manuscript.*

Abbreviations

The following abbreviations are used in the text.

CD	Circular dichroism
E. coli	Escherichia coli
k_{ass}	Second order association rate constant
K_d	Equilibrium dissociation constant
k_{diss}	Dissociation rate constant
K_i	Inhibition constant
k_{obs}	Observed pseudo-first-order rate constant
NMR	Nuclear magnetic resonance

Introduction

Enzymes are responsible for catalyzing practically all chemical reactions required for the existence of life. For chemical reactions to occur, the reactants must overcome the activation energy barrier. Enzymes assist the reaction course by bringing the reactants together in the right geometrical orientation and by stabilizing the transition state intermediates, thereby lowering the activation energy barrier and increasing the rate of the reaction. An important group of enzymes are those that cleave the peptide bonds that link amino acid residues in the polypeptide chains of peptides and proteins. These enzymes are called peptidases (or proteinases to emphazise the protein nature of the substrate), and are classified as exo- or endo- peptidases depending on whether they cleave their substrates from one of the ends or inside the polypeptide chain, respectively. Another way of classifying proteinases is by referring to the mechanism they use to hydrolyse the peptide bond. According to the latter view, proteinases are divided into four classes, serine, cysteine, aspartate and metallo-proteinases. As proteinases are potentially harmful for the organism, the activity of endogenous enzymes is tightly controlled at different levels, e.g. by the regulation of transcription levels of the mRNA coding for these enzymes, by the expression of inactive pro-enzymes, by the localization the active enzymes in isolated compartments and by the presence potent inhibitors where proteinase activity is not desirable.

The aim of this thesis has been to study the mechanism by which such an inhibitor, cystatin A, in mammalian organisms exerts its inhibition of exogenous and endogenous cysteine proteinases of the papain family. This inhibitor was cloned and expressed in *E. coli*, and the inhibitory properties of the wild-type cystatin A and of several variants of the inhibitor towards a number of target proteinases were analysed.

Previous investigations

Cysteine proteinases of the papain family

According to the latest classification there are 42 families of cysteine proteinases (Rawlings and Barrett, 1998), of which the papain superfamily is the most abundant and best characterized. Other well-studied cysteine proteinase families are the calpain family, including the calcium-dependent cytosolic enzyme calpain, known from mammals and birds, and the streptopain family, a group of cysteine proteinases found in *Streptococcus*.

Cysteine proteinases of the papain family are widely distributed and are found in species such as baculovirus (Rawlings et al., 1992), yeast (Enenkel et al., 1993), protozoa (Scholze and Tannish, 1994; North, 1992), and in plants and animals (reviewed in Rawlings and Barrett, 1994).

Papain, a plant enzyme found in the latex and vacuoles of papaya, was the first cysteine proteinase to be isolated and characterized (Balls et al., 1937) and has therefore given name to the superfamily. Other known papain-like enzymes of plant origin are chymopapain and caricain, also from papaya, actinidin from kiwi fruit, ficin from fig and bromelain from pineapple. The function of these plant proteinases is largely unknown, but enzymes localized in the in the vacuoles of the fruit are believed to protect against predators. Moreover, the presence of the proteinases in the latex presumably promotes the coagulation of the latter (Boller, 1986).

In mammals, papain-like proteinases are in general confined to the lysosomes, where most of the cellular protein catabolism is effected by an array of endo- and exopeptidases (Brocklehurst et al., 1987); Bohley and Seglen, 1992; Kirschke et al., 1995). Lysosomal proteinases have been given the generic name cathepsins, proposed by Willstätter and Bamann (1929) for proteinases which, like most mammalian papain-like cysteine proteinases, display optimum activity at acidic pH. Cathepsins B, C, H, L, K and S are lysosomal cysteine proteinases, of which cathepsins B, C, H and L are ubiquitous in mammalian cells, whereas cathepsin K and S have more specific distribution (Barrett and Kirschke, 1981; Drake et al., 1996; Kirschke and Wiederanders, 1994). Cathepsin K is predominantly expressed in osteoclasts, whereas cathepsin S is unevenly distributed between organs, with a high occurrence in tissues such as spleen and lung, and exhibits high endopeptidase activity even at neutral pH (Kirschke et al., 1989; Quian et al., 1989).

Biosynthesis of lysosomal cysteine proteinases

The lysosomal cysteine proteinases are synthesized as inactive pre-proenzymes and are subject of several post-translational modifications before they reach the lysosomes as mature enzymes. These modifications include cleavage of the pre-fragment during translocation to the Golgi network, where disulphide bonds are formed and glycosylation takes place, followed by phosphorylation of the attached mannoses. This last step targets the proenzyme to the lysosomal compartment via the mannose-6-phosphate receptor. Inside the lysosomes, with an acidic environment, the enzyme is activated by cleavage of its propeptide, which blocks the active site cleft in the pro-form (reviewed in Kirschke et al., 1995). The lysosomal cysteine proteinases are thus evidently tightly controlled only to be active inside the lysosomes. This is further evidenced by the irreversible inactivation of the mature enzymes at neutral pH (Barrett and Kirschke, 1981; Turk et al., 1995a), and by the occurrence of

potent protein inhibitors of these enzymes in the cytosol of the cells, as well as extracellularly (see below).

Structure and mechanism of action of cysteine proteinases

The majority of the enzymes belonging to the papain family are relatively small proteins active as monomers, with molecular masses ranging from 20 to 35 kD. Similarities in the primary and three-dimensional structures of enzymes for which structural data are available indicate that they have evolved from a common ancestor (Brocklehurst et al., 1987; Rawlings and Barrett, 1994; Kamphuis et al., 1995; Turk et al., 1997). The homology is particularly evident by the high degree of conservation of the amino acid sequence around the catalytically active residues Cys-25 and His-159 (papain numbering will be used throughout) and also around Asn-175, which is conserved in nearly all family members (Rawlings and Barrett, 1994).

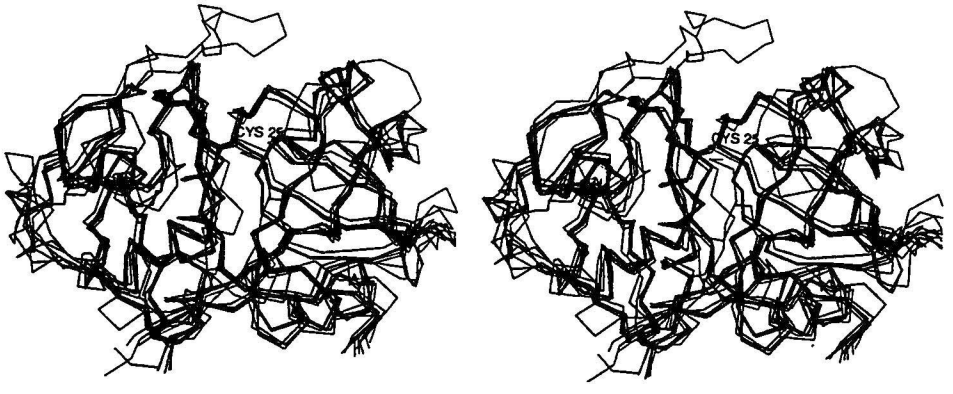

Fig. 1. Structural similarities of the papain-like cysteine proteinases. Plot of the backbones of ten enzymes are superimposed on each other. Structures are shown in the standard view along the two-domain interface with the active-site cleft at the top. (Reprinted, with permission, from Turk et al. (1998). Copyright 1998, Walter de Gruyter & Co, Berlin-New York.)

Not surprisingly, it was the three-dimensional structure of papain which was the first to be resolved (Drenth et al., 1976; Kamphius et al., 1984). Subsequently, several other structures of papain-like cysteine proteinases have been elucidated by X-ray crystallography. Despite these proteinases having widely different origins, e.g. from plants, animals and parasites, their three-dimensional structures reveal the same common fold (Kamphuis, 1995; Turk et al, 1998). This is exemplified in Figure 1, where the backbones of cruzain (the catalytic domain of cruzipain from *Trypanosoma cruzi*) (McGrath et al., 1995), procathepsin L (Coulombe et al., 1996), glycyl endopeptidase (O'Hara et al., 1995), human cathepsin B (Musil et al., 1991), caricain (Pickersgill et al.,

1991), chymopapain (Maes et al., 1996), actinidin (Baker, 1980), porcine cathepsin H (Guncar et al., 1998) and papain (Kamphuis et al., 1984) are superimposed on each other.

A prominent feature of the tertiary structure of the papain-like cysteine proteinases (see Figure 1, which shows the "standard" view), is that the enzymes fold into two main domains separated by a V-shaped cleft, the active site cleft. The two residues participating in the peptide cleavage, Cys-25 and His-159, are located one in each domain, pointing to the inside of the cleft. The substrate binds along the bottom of the active site cleft, making a number of contacts with the proteinase. It is commonly accepted that the enzyme can interact with as many as seven residues from the substrate (Schechter and Berger, 1967; Del Nery et al., 1997). These residues are denoted P_4-P_1 and P_1'-P_3', the cleavage occurring between P_1 and P_1', and the corresponding binding places for them in the enzyme are called subsites; S_4-S_1 and S_1'-S_3', respectively (Figure 2). This substrate binding model has recently been questioned by Turk et al. (1998), after comparing structural data from several papain-like cysteine proteinases. Instead, their conclusion is that only subsites S_3-S_1 and S_1'-S_2' are well defined and that any additional interactions between the enzymes and residues of the substrate do not arise from defined binding pockets common for all enzymes. Lysosomal cysteine proteinases display a broad substrate specificity, in accordance with their proposed "housekeeping" functions (see below), but have preferences regarding which side-chain of the substrate will be accommodated in the S_2 subsite. Cathepsins L and S have quite similar specificities for substrate residues in position P_2, preferring large hydrophobic or aromatic side chains, such as Phe or Val, whereas cathepsin B also accepts positively charged residues, such as Arg, in that position (Barrett and Kirschke, 1981; Kirschke et al., 1995).

Most of the lysosomal cysteine proteinases are endopeptidases, with the exception of cathepsin C, which is an aminopeptidase. Cathepsin B and cathepsin H display both endo- and exopeptidase activity. The carboxypeptidase activity of cathepsin B can be explained by its structure, which reveals that a loop region of about 20 residues, called the occluding loop, partially blocks free access of protein substrates to the active site cleft of the enzyme (Musil et al., 1991). By the removal of this loop in a recombinant cathepsin B variant the enzyme completely lost exopeptidase activity (Illy et al., 1997). In the case of cathepsin H, part of the propeptide, called the mini-chain, remains attached to the mature enzyme by a disulphide bond. This mini-chain blocks the unprimed substrate binding sites of the enzyme up to the S_2 subsite (Guncar et al., 1998) and thus explains the aminopeptidase activity of cathepsin H.

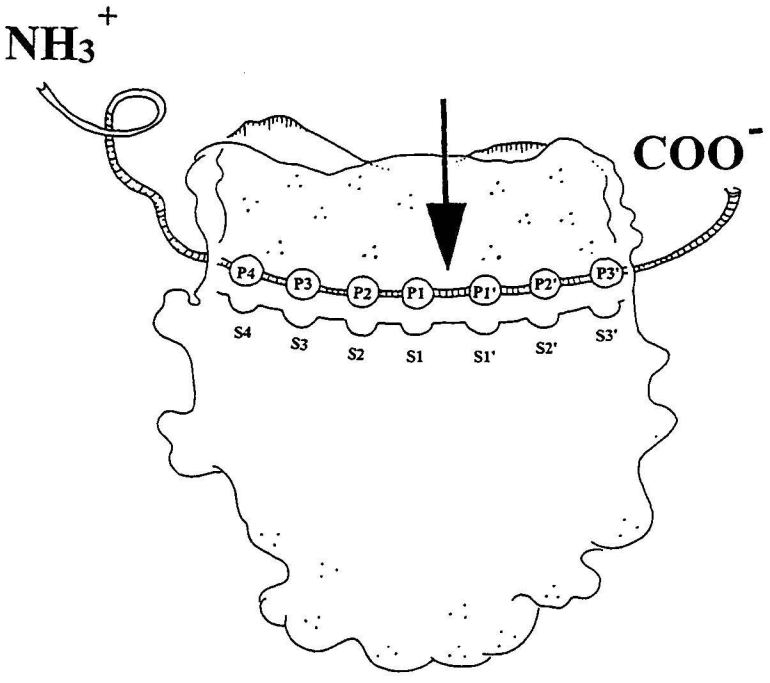

Fig. 2. Schematic representation of a protein substrate bound to a papain-like cysteine proteinase. The residues of the substrate interacting with the active-site cleft of the enzyme are denoted P_4 to P_1 and P_1' to P_3', the scissile bond being between P_1 and P_1' (marked with an arrow). The corresponding substrate-binding subsites of the enzyme are S_4 to S_1 and S_1' to S_3'.

The detailed mechanism by which the archetypal enzyme papain catalyses the hydrolysis of a peptide bond is not fully elucidated at the atomic level (Storer and Menard, 1994). It is, however, proven that the thiolate-imidazonium ion pair formed by Cys-25 and His-159 is involved. The basic steps of the mechanism, after the initial binding of the substrate to the active site cleft of papain, are presented in Figure 3. First, a nucleophilic attack of the thiolate ion on the carbonyl carbon of the scissile bond to the P_1 residue of the substrate leads to the formation of a thiolester between Cys-25 and the carbonyl atom of the P1 residue (Step 1a). In this first step, the imidazonium ion of His-159 donates a H^+ to the N-terminal amine of the newly cleaved peptide, which was the P1' residue of the original substrate, and this generated peptide diffuses from the active site of the enzyme (Step 1b). The second step involves nucleophilic attack by a water molecule on the carbonyl atom in the thiolester bond between Cys-25 of the enzyme and the remaining segment of the original peptide (Step 2a). The consequence of this nucleophilic attack is the addition of a hydroxyl group to the carbonyl carbon of the P1 residue, and the subsequent break of the thiolester bond, which leaves the other part of the cleaved peptide free to diffuse from the active site cleft of the enzyme (Step

2b). In this step, His-159 facilitates the nucleophilic attack by accepting a H^+ from the water molecule. The two briefly described steps are called the acylation and the deacylation steps (Storer and Menard, 1994) and presumably involve formation of transient tetrahedral intermediates, in analogy with the mechanism of hydrolysis of serine proteinases, which is known in greater detail.

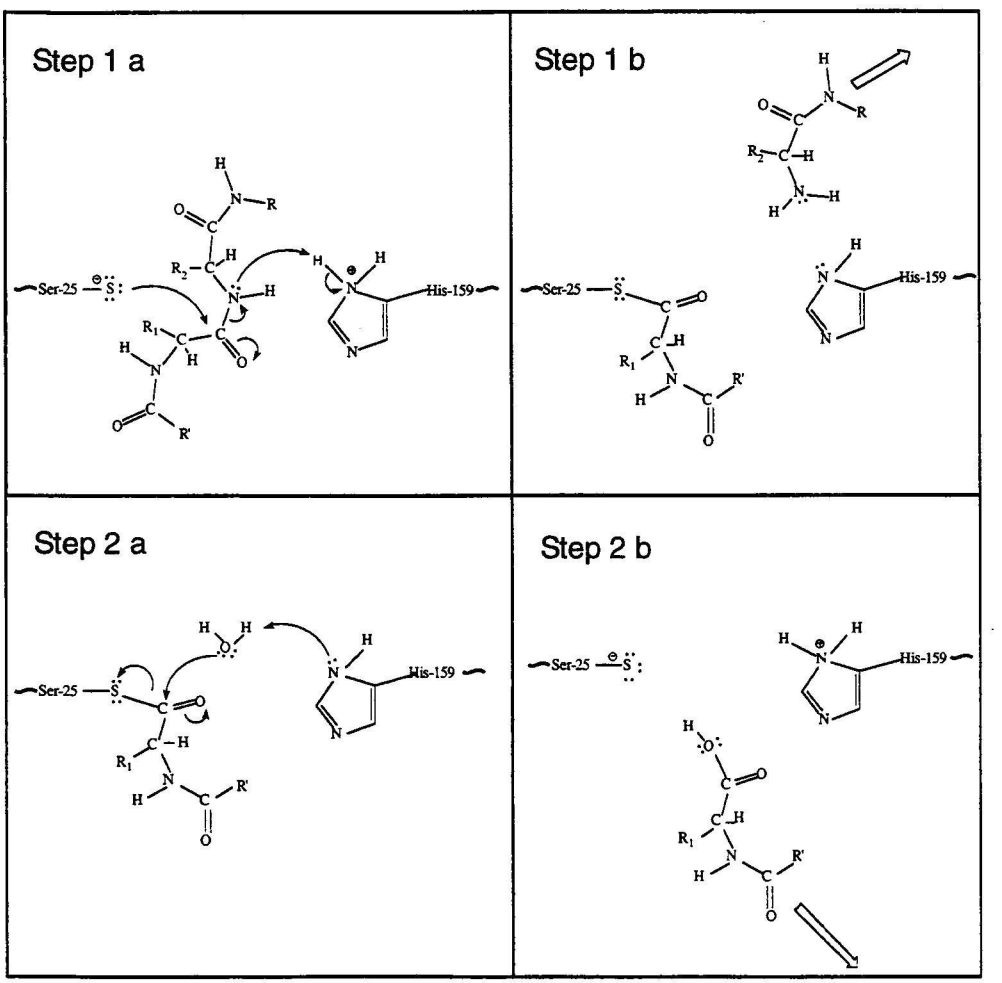

Fig. 3. Two basic steps involved in the hydrolysis of a peptide bond by the cysteine proteinase papain. Step 1 is the acylation step, in which the scissile bond is cleaved and a thiolester bond is formed between Cys-25 of papain and the substrate. Step 2 is the deacylation step, in which the cleaved product is released and the catalytically active thiolate-imidazonium ion is regenerated.

Functions of lysosomal cysteine proteinases

The physiologically most important role of lysosomal cysteine proteinases is the degradation of proteins imported into lysosomes, as mentioned earlier. It has been estimated that up to 80% of the peptidase activity in the mammalian organism takes place inside the lysosomes, and that the cysteine proteinases are the most active among the lysosomal proteolytic enzymes (Barrett and Kirschke, 1981). Cathepsin B is the most abundant, but cathepsin L and cathepsin S exhibit the highest activities among the papain-like proteinases (Barrett and Kirschke, 1981; Mason et al., 1985; Kirschke et al., 1989). A recent report, however, claims that the concentration of both cystatins B and L can be equally high, up to 1 mM, inside the lysosomes of cultured cell lines (Xing et al., 1998). Besides the general degradative properties of lysosomal cysteine proteinases, they have been attributed other functions. Cathepsin B has been proposed to be involved in protein degradation for antigen presentation (Mizuochi et al., 1994; Authier et al., 1996), and cathepsin K is implicated in bone remodelling (Drake et al., 1996; Inui et al., 1997). Moreover, it has been postulated that cathepsin B is involved in the processing of pro-forms of peptide hormones (Neves et al., 1996; Dunn et al., 1996).

Pathological conditions related to cysteine proteinase activity

Due to the broad substrate specificity, the wide distribution and the high specific activity of the papain-like cathepsins it is not surprising that they have been implicated in the tissue destruction occurring under various pathological conditions, such as arthritis (Iwata et al., 1997; Lemaire et al. 1997), Alzheimer's disease (Mantle et al., 1995; Mackay et al., 1997) and osteoporosis (Bossard et al., 1996). Interestingly, it was found that deficiency of the osteoclast-specific enzyme, cathepsin K, causes osteosclerosis and short stature in an autosomal recessive disease called pycnodysostosis (Gelb et al., 1996).

Many indications that implicate cathepsin B, and to a lesser extent cathepsin L, in cancer progression have attracted special interest to the role played by the two cathepsins in this disease. Increased levels of mRNA coding for cathepsin B and cathepsin L, as well as higher concentrations of the expressed enzymes, have been reported in several types of human tumors, including breast tumors (Sloane et al., 1987; Rozhin et al., 1989; Chauhan et al., 1991). Recently, it was also shown that invasive breast tumor cell lines contained higher concentrations of active cathepsins B and L than did similar, but non-invasive, cells (Xing et al., 1998). In addition, several investigations have found increased levels of cathepsin B in lung tumors. The enzymatic activity of cathepsin B was showed to be higher in lung tissue from all patients that suffered from primary lung adenocarcinoma when compared with that of healthy lung tissue (Krepela et al., 1990). Also, Werle et al. (1994) noted that

lung tumors of different cell types were considerably richer in cathepsin B activity than surrounding, normal tissue. Furthermore, in brain tumors of the glioblastoma type, transcription of cathepsin B was found to be 8 to 10 fold higher than in normal brain (Sivaparvathi et al., 1995), and Rempel et al. (1994) reported that this enhanced transcription was correlated with the invasiveness of the tumors. It has been shown *in vitro* that cathepsins B, K and L are able to degrade laminin, fibronectin and collagen IV, which are components of the basement membrane, at physiological pH (Lah et al., 1989; Buck et al., 1992; Kafienah et al., 1998). Furthermore, Higashiyama et al. (1993) and Inoue et al. (1994) have reported positive staining for cathepsin B correlated with the degradation of laminin in lung adenocarcinoma. One mechanism by which cathepsin B may promote tumor progression is thus by directly degrading the laminin and collagen IV of the basement membrane (Yan et al. 1998).

Protein inhibitors of papain-like cysteine proteinases

In the search for proteins capable of inhibiting papain-like cysteine proteinases Fossum and Whitaker (1968) purified and partially characterized a protein from chicken egg-white, which displayed potent inhibitory activity towards papain and ficin (Fossum and Whitaker, 1968; Sen and Whitaker, 1973). Other studies established the ability of this protein to inhibit all members of the papain family that were known to that date, e.g. cathepsins B, C, H and L (Keilova and Tomasek, 1977; Anastasi et al., 1983). It took more than a decade before the primary structure of this protein, named cystatin by Barrett (1981), was determined (Turk et al., 1983; Schwabe et al., 1984). Meanwhile, reports characterizing new cystatins from other sources were emerging in the literature. Several human inhibitors were described, such as cystatin A from epidermis, (Järvinen, 1978), cystatin B from spleen and liver (Järvinen and Rinne, 1982; Green et al., 1984), cystatin C (or γ-trace) from urine and cerebrospinal fluid (Grubb and Löfberg, 1982) and cystatins S, SA and SN from saliva and tears (Isamura et al., 1984, 1986, 1987). Early discoveries of mammalian, non-human cystatins included rat cystatins α and β (Takeda et al., 1983; Takio et al., 1983, 1984; Hirado et al., 1981).

As sequence data began to accumulate, classification of the cystatins on the basis of their sequence homology became possible, and these inhibitors were grouped into three main subfamilies, all belonging to the cystatin superfamily.

Family I

Inhibitors from family I consist of a single, non-glycosylated polypeptide chain of about 100 amino acids, having a molecular mass in the range of 11-12 kD. They lack internal disulphide bonds, and their localization is primarily

intracellular. The only human cystatins of family I that are known to date are cystatin A and B (also called stefin A and B). There is a different pattern in the distribution of these inhibitors in the organism; whereas cystatin B is almost ubiquitously expressed in all cells, the occurrence of cystatin A is more restricted, e.g. to epidermis, leukocytes and liver (Järvinen et al 1987; Brzin et al., 1983; Green et al., 1984). Corresponding cystatins from other mammals have also been isolated and characterized, including cystatin A and B from mouse (Tsui et al., 1993), rat (Järvinen, 1976; Takio et al., 1983, 1984) and cattle (Krizaj et al., 1992; Turk et al., 1995b), all showing a high degree of homology over the species borders. Lately, the discovery of bovine stefin C (Turk et al., 1995b) and pig stefin D (Lenarcic et al., 1996) and have added new members to family I.

Family II

The inhibitors of family II are also single-chain proteins, like cystatins from family I. However they are slightly larger, consisting of about 115 residues, with molecular masses in the range of 12-14 kD, and contain two internal disulphide bonds. Most of them are non-glycosylated, although N-glycosylation has been reported for some members (Ni et al., 1997; Sotiropoulou et al., 1997). Cystatins of family II are expressed with a signal peptide, which targets the proteins to extracellular compartments; hence these cystatins are found in body fluids. The best studied member, human cystatin C, has been reported to be present in all fluids examined (Abrahamson et al., 1986). This widespread distribution is also shared by cystatin M, which was recently identified (Sotiropoulou et al., 1997). Cystatins S, SA and SN are found predominantly in saliva and tears (Isamura et al., 1984, 1986, 1987) and cystatin D only in saliva (Freje et al., 1991), whereas the tissue specificity of cystatin E appears to be restricted to the uterus and the liver (Ni et al., 1997). Inhibitors homologous to human cystatin C have been found in cattle (Hirado et al., 1985; Olsson et al., 1997), mouse (Solem et al., 1990), rat (Esnard et al., 1988), dog (Poulik et al., 1981), sheep (Tu et al., 1990) and monkey (Grubb et al., 1982). The avian counterpart of cystatin C, chicken cystatin, has also been extensively studied, and was the first cystatin for which the crystal structure was elucidated (Bode et al., 1988). Chicken cystatin has the advantage of being relatively simple to purify from chicken egg-white (Anastasi et al., 1983), and certainly, before the introduction of recombinant protein techniques, this property played a major role for the choice of inhibitor to study. Also cystatins found in the venom of two snakes, African puff adder and Taiwan cobra, and in the large granules of hemocytes of the Japanese horseshoe crab belong to family II (Evans and Barrett, 1987; Agarwala et al., 1996; Brillard-Bourdet et al., 1998). Of the mammalian cystatins, it is cathepsin M which shows highest homology with these non-mammalian cystatins (Brillard-Bourdet et al., 1998).

Family III

The kininogens are members of the third cystatin superfamily. These family III cystatins are glycosylated proteins present in blood, with molecular masses ranging from 50 to 90 kD (Turk et al., 1995c). The kininogens contain three domains homologous to cystatins of family II. Two of the domains are able to inhibit papain-like cysteine proteinases (Salvesen et al., 1986; Turk et al., 1995c), and one of these confers to the kininogens their calpain-inhibiting activity (Salvesen et al., 1986). Three different kininogens have been described, low- and high-molecular weight kininogen from humans and many other mammals, which are splice variants of the same gene (Müller-Esterl et al., 1986), and rat T-kininogen (Esnard and Gauthier, 1983). Beside being inhibitors of cysteine proteinases from the papain and calpain families, kininogens also have other physiological activities. As indicated by their name, kininogens were first recognized as precursors of kinins, which are vasoactive peptides released from the mature single polypeptide chain of kininogen by the action of kallikreins (Müller-Esterl et al., 1986). Moreover, high-molecular-weight kininogen participates in the contact phase of blood clotting (DeLa Cadena and Colman, 1991). A recent report also suggests that kininogens are ferritin-binding proteins (Torti and Torti, 1998).

Other cystatins

An increasing number of cystatins that do not unambiguously belong to any of the three classical families have been identified in animals other than mammals and also in plants. These cystatins are all inhibitors of papain-like cysteine proteinases but are not sufficiently homologous with members of the classical families to be classified as belonging to any of these. The most studied of these inhibitors are oryzacystatin I and II, found in rice and resembling cystatins from family I (Abe et al., 1988; Kondo et al., 1990). Other plant cystatins are found in corn (Abe et al., 1992), African berry (Murzin, 1993), papaya (Song et al., 1995), avocado (Kimura et al., 1995) and soybean (Misaka et al., 1996). Furthermore, cystatin-like inhibitors have been found in fish (Koide and Noso, 1994; Tsai et al., 1996), in parasitic protozoa (Irvine et al., 1992) and in bacteria (Tsushima et al., 1992).

Physiological role

The general belief is that the primary physiological function of cystatins is to prevent uncontrolled proteinase activity outside the lysosomes, both intracellularly and extracellularly. However, no direct evidence for the importance of these inhibitors in controlling peptidase activity under normal physiological conditions is available. Instead, the protective role of cystatins is

indicated by increased levels of cystatins in locations where elevated proteinase activity is observed (reviewed in Henskens et al., 1996). Such increased levels of cystatins have been observed in periodontal inflammations (Henskens et al 1993, 1994, 1996; Naito et al., 1992; Alves et al., 1994), skin inflammations (Hopsu-Havu et al., 1983; Järvinen et al., 1987) and in rheumatoid arthritis (Lenarcic et al., 1988). Besides, it has been demonstrated that cystatin C inhibits bone resorption in vitro, probably by inhibiting the activity of osteoclastic proteolytic enzymes released into the resorption lacunae (Lerner et al., 1997). On the other hand, also the inhibitory capacity of cystatins is affected by the action of proteinases. For instance, the N-terminal segment of cystatin C is cleaved by neutrophil elastase and by bacterial proteinases during periodontitis (Abrahamson et al., 1991, 1997; Grenier, 1996), which greatly diminishes the affinity of the inhibitor for target proteinases. Cystatin A has also been shown to be the subject of inactivation by cleavage of its N-terminal region by an aspartic proteinase from the fungus, *Candida albicans* (Tsushima et al., 1994).

Special emphasis has been put on elucidating the role of cystatins in tumor progression. Decreased levels of cystatin A were found in breast cancer (Lah et al., 1992), and Knoch et al. (1994) reported an imbalance between the levels and activity of cathepsin B and those of cystatin B in lung cancer. Recently, the expression of cystatin M was shown to be suppressed in human breast cancer (Sotiropoulou et al., 1997), as was found for cystatin B in oesophagus carcinoma (Shiraishi et al., 1998). Moreover, the inhibitory efficiency of cystatin A from human sarcoma was decreased (Lah et al., 1989). The lowered levels and activities of cystatins, in correlation with increased proteinase activity, thus seems to facilitate tumor progression.

Besides inhibiting endogenous cysteine proteinases, cystatins most probably are involved in the defence against invading microorganisms by inactivating foreign proteinases used by these invaders to degrade the host organism. Cystatins have thus been shown to inhibit cysteine proteinases from parasites (Stoka et al., 1995; Chagas et al., 1997; Troeberg et al., 1996). Furthermore, cystatin C and cystatin D were reported to inhibit corona virus replication at physiological concentrations (Collins and Grubb, 1998), and replication of herpes simplex virus was inhibited by cystatin C (Björck et al., 1990) and oryzacystatin (Aoki et al., 1995). There are also reports of cystatins inhibiting bacterial growth, as in the case of *Porphyromonas gingivalis* (a bacterial strain involved in periodontitis), the growth of which is inhibited by cystatin S and chicken cystatin (Naito et al., 1995; Umemoto et al., 1996; Blankenboorde et al., 1996).

Diseases caused by anomalous cystatins

Two pathological conditions in which cystatins are directly involved have attracted special attention. One of these is a hereditary form of amyloid angiopathy found in Icelandic families. This autosomal dominant disorder leads to brain haemorrhage due to amyloid deposition in the blood vessels of the brain, and eventually to death in affected patients, generally before the age of 40 years (Jensson et al., 1987). The affected gene for the cystatin associated with this disease, cystatin C, has a single base mutation, resulting in a Leu-68 to Gln replacement in the protein (Palsdottir et al., 1988; Abrahamson et al., 1992). The mutated cystatin C has been shown to dimerize and form complexes (Abrahamson and Grubb, 1994), which presumably is the explanation why it is deposited in amyloid fibrils. The other disorder, which is a rare autosomal recessive form of myoclonus epilepsy, has been traced to mutations in the cystatin B gene (Pennacchio et al., 1996). The most common of these mutations is an expansion of a minisatellite repeat located upstream of the transcription start of the gene (Lalioti et al., 1997b; Virtaneva et al., 1997; Lafreniere et al., 1997), causing decreased levels of cystatin B mRNA. Interestingly, one of the point mutations observed in a patient affected by this disorder was a G to C change in exon 1, which results in the replacement of the evolutionarily conserved Gly-4 of cystatin B by Arg (Lalioti et al., 1997a), thereby almost abolishing the inhibitory activity of the protein towards target cysteine proteinases (see Present investigation). The underlying mechanism by which the absence of cystatin B activity leads to the disease is as yet unknown.

Structure of cystatins and general mechanism of inhibition

By the middle of the 1980's several general aspects of the interactions of cystatins with their target proteinases had been ascertained. From studies by Turk et al. (1986), Barrett (1986) and Green et al. (1984), and many others, it was known that cystatins bound equimolarly and reversibly to their target enzymes, displaying very high affinity, with inhibition constants as low as < 5 pM. Homology alignments of those amino acid sequences of cystatins that had been determined gave the first hints on which regions of the inhibitors could be involved in the binding to the target proteinases. The attention was attracted by the evolutionarily most conserved residues, which are a Gly close to the N-terminus of the inhibitor (Gly-9 in chicken cystatin) and a segment with the consensus sequence Gln-Val-Val-Ala-Gly of almost fully conserved residues in the central part of the protein. Several experimental results that had been obtained by this time also supported the involvement of these regions in proteinase binding. Abrahamson et al. (1987) thus reported that cystatin C and chicken cystatin whose N-terminal region had been deleted before the conserved Gly residue lost affinity for papain by more than three orders of

magnitude. Takeda et al. (1985) and Isemura et al. (1986) made similar observations when investigating the properties of rat cystatin A and cystatin S lacking the N-terminal segment, in which the conserved Gly is located. Contradicting these findings, there was a report of a plant cystatin, oryzacystatin I, that maintained activity despite truncation of 21 of its most N-terminal amino acids (Abe et al., 1988). Evidence for the importance of the central QVVAG region for proteinase binding was provided by the observation that, of the three cystatin-like domains of kininogen, only those having the QVVAG sequence were able to inhibit papain and cathepsin L (Salvesen et al., 1986). As the two presumed binding regions of the inhibitor may have been required for proper folding and not for direct interaction with the proteinases, synthetic peptides containing amino acids from these regions were tested for inhibitory activity. A peptide comprising residues 4-21 of cystatin C showed no inhibition of papain (Abrahamson et al., 1987), in contrast with a synthetic peptide that harboured the central QVVAG segment of the cystatins, which inhibited the activity of papain and cathepsin B to some extent (Teno et al., 1988).

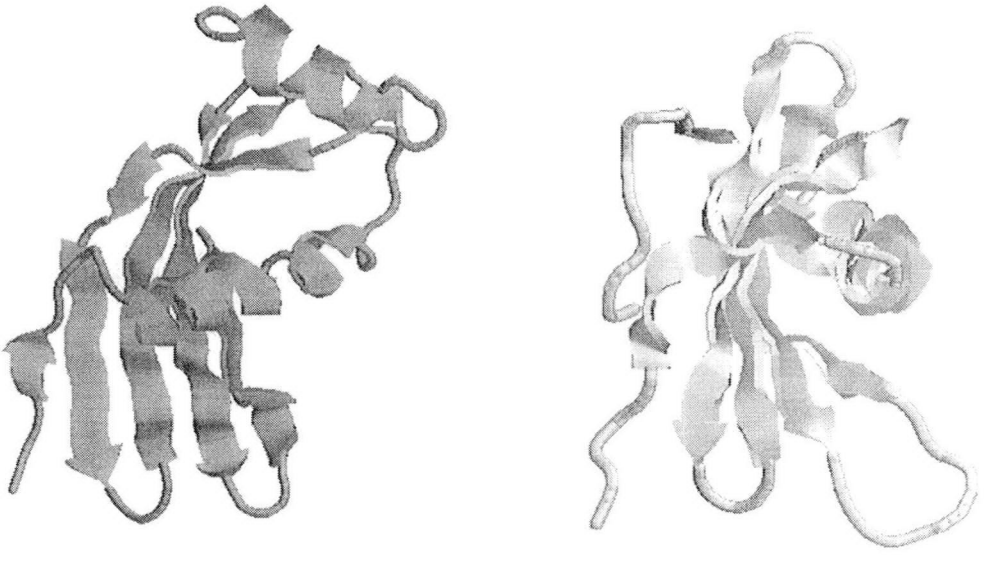

Fig. 4. Crystal structure of chicken cystatin (Bode et al., 1988) (left) and solution structure of cystatin A (Martin et al., 1995) (right). The proteins are represented as cartoons with the the regions that form the inhibitory wedge facing down.

The elucidation of the crystal structure of chicken cystatin (Bode et al., 1988), unfortunately of a form lacking the amino acids preceding the conserved

Gly-9, provided strong support for the previously implicated regions of the inhibitor participating in the binding to target enzymes. The crystallized cystatin molecule consists mainly of a five-stranded antiparallel β-pleated sheet wrapped around a five-turn α-helix and a short appending α-helix segment (Figure 4). This three-dimensional structure of chicken cystatin shows that the N-terminal Gly-9 residue, a central hairpin loop containing the QVVAG sequence (QLVSG in chicken cystatin), and a C-terminal loop, with the evolutionarily rather conserved Pro-103 and Trp-104 (in chicken cystatin numbering), are all exposed on the surface of the protein, forming a wedge-shaped edge (Figure 4). On basis of computer docking experiments of the cystatin and papain structures, it was suggested by Bode et al. (1988) that this edge is the region of the inhibitor that binds to the active site cleft of papain, thereby blocking the access of substrates to the latter and, consequently, inhibiting the enzyme. The docking experiments also indicated that the two interacting regions, the "inhibitory wedge" of the cystatin and the active-site cleft of papain, are highly complementary and that thus no major conformational changes are needed on either of the proteins for high affinity binding.

The proposed docking model was essentially corroborated by the crystal structure of a complex between cystatin B, a cystatin from family I, and papain reported two years later (Stubbs et al., 1990). Briefly, the structure of the complex showed that the two residues of the inhibitor immediately adjacent to the conserved Gly-4, *viz.* Ser-3 and Met-2, were positioned in a substrate-like manner, interacting with the putative S_2 and S_3 subsites of papain, respectively. This positioning of the N-terminal residues created a type II turn in the backbone of the inhibitor, only permitted by a Gly residue in position 4, explaining the absolute conservation of this residue in cystatins. Furthermore, the first and second binding loops were indeed very complementary to the active site groove of papain, as proposed by Bode et al. (1988), creating many rather unspecific and mostly hydrophobic interactions between the Gln-46–Asn-52 and Gln-71–His-75 regions of the inhibitor and the groove of the enzyme. In addition, the C-terminal segment of cystatin B also contributed to the binding of papain by its Tyr-97 and Phe-98 residues.

More recently, NMR studies have provided the three-dimensional solution structures of chicken cystatin (Dieckmann et al., 1993) and cystatin A (Martin et al., 1995; Tate et al., 1995; Figure 4), and given indications on the secondary and tertiary structure of cystatin C (Ekiel et al., 1997). All these cystatin structures show very similar folding, especially in the loop regions and around the N-terminal Gly residue, indicating that the model of binding to papain suggested for chicken cystatin (Bode et al., 1988) and observed for cystatin B (Stubbs et al., 1990) probably is valid for most cystatins. The solution structures also provided additional information, such as the finding that the N-terminal region possesses a high degree of flexibility, whereas the second binding loop is less flexible, and the first binding loop is very rigid.

Moreover, the short appending helix that is present in the crystal structure of chicken cystatin is not observed in the corresponding solution structure, perhaps only being an artifact of the crystal packing.

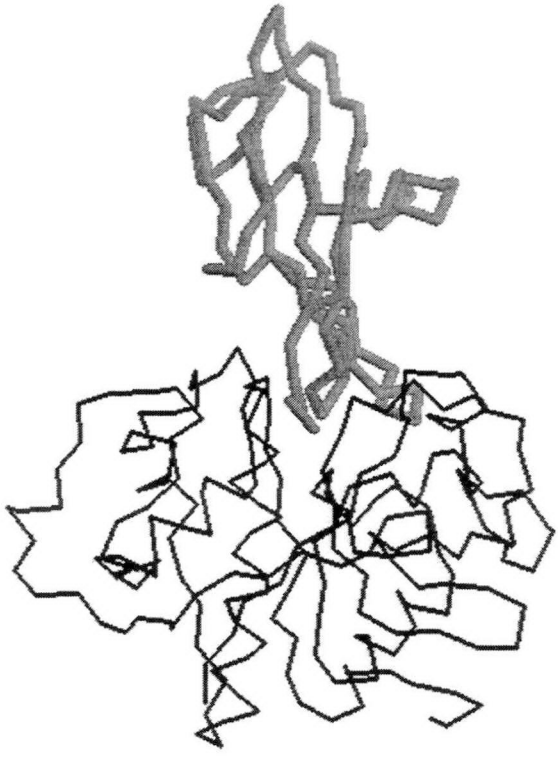

Fig. 5. Stucture of the complex between cystatin B and papain (Stubbs et al., 1990). The inhibitor is on top (thick lines). The complex is shown in a view along the active-site cleft of papain, with the N-terminal region of cystatin B at the front.

Another clue to the mechanism by which cystatins inhibit papain-like cysteine proteinases was provided by stopped-flow kinetics studies of the association of chicken cystatin and papain at high inhibitor concentrations (Björk et al., 1989). It was shown that, under pseudo-first order conditions (i.e. at an excess of inhibitor over enzyme), the observed rate constant by which the inhibitor bound to the enzyme increased linearly with increasing inhibitor concentration. Moreover, the high second-order rate constant, around 10^7 $M^{-1} \cdot s^{-1}$, indicated that the reaction was close to diffusion-controlled. These results presented strong kinetic evidence for a one-step inhibition mechanism, as described in Scheme 1, supporting the model proposed from docking studies, in which a good fit between chicken cystatin and papain was achieved without conformational changes (Bode et al., 1988).

$$E + I \underset{k_{diss}}{\overset{k_{ass}}{\rightleftarrows}} EI \qquad \text{(Scheme 1)}$$

In addition, by separate measurements of the association and dissociation rate constants for the interactions between inhibitor and enzyme, it was possible to accurately determine the extremely high affinity of chicken cystatin for papain. These analyses gave a K_i value of ~ 60 fM, whereas only a limit for K_i of < 5 pM had been estimated by equilibrium methods (Green et al., 1984), i.e. a ~100-fold higher affinity could be determined by this procedure. Most subsequent studies of the association kinetics between cystatins and papain-like cysteine proteinases, conducted under pseudo-first order conditions and at high inhibitor concentrations, are consistent with this mechanism. This includes interactions between chicken cystatin and the proteinases ficin, chymopapain A or actinidin (Björk and Ylinenjärvi, 1990), between cystatin C and papain or actinidin (Lindahl et al., 1992a) and between bovine stefin B and cathepsin S (Turk et al., 1994). Moreover, these studies showed that the affinities of the cystatins tested towards papain, ficin and cathepsin S all are very strong, with K_i values between 0.01 and 8 pM, whereas actinidin is inhibited by cystatin C or chicken cystatin more than three orders of magnitude more weakly than these enzymes, although still with appreciable affinity. In contrast with these findings, the interaction between cystatin C and cathepsin B has been shown by similar methods to occur via a two-step mechanism (Nycander et al., 1998) (Scheme 2). This behaviour presumably reflects the need for the inhibitor to displace the occluding loop of cathepsin B (Musil et al., 1991) before gaining access to the partially covered active site cleft of the enzyme. Due to this mechanism, cathepsin B is inhibited more weakly ($K_i \geq 0.2$ nM) by cystatin C, and other cystatins, than most cysteine proteinases.

$$E + I \overset{K_1}{\rightleftarrows} EI \underset{k_{-2}}{\overset{k_{+2}}{\rightleftarrows}} EI^* \qquad \text{(Scheme 2)}$$

Chemically and genetically engineered variants of cystatins

In order to assess the importance of the different regions forming the "three-partite" wedge-shaped edge of the cystatins for proteinase binding, several studies have investigated how modifications of these regions affect the affinity and kinetics of binding of the inhibitors to target proteinases. N-terminally truncated variants of rat cystatin A (Takeda et al., 1985), chicken cystatin (Abrahamson, et al., 1987; Machleidt et al., 1989; Lindahl et al., 1992b) and cystatin C (Abrahamson, et al., 1987, 1991; Hall et al., 1995, 1998; Björk et

al., 1994, 1996) were reported to have appreciably decreased affinities for papain-like cysteine proteinases. In discrepancy with these results were the reports that oryzacystatin I, as already mentioned, and human stefin B maintained their inhibitory activity on truncation of 21 and 6 of their N-terminal residues, respectively (Abe et al., 1988; Thiele et al., 1990). However, the former results have been disputed by Urwin et al. (1995) in a study in which deletion of the 21 most N-terminal residues of oryzacystatin I, and mutations of the conserved N-terminal Gly residue, resulted in the loss of inhibitory ability.

The N-terminal region of chicken cystatin was estimated to contribute about 40% of the total binding energy to the complex formation with papain and actinidin (Lindahl et al., 1992b), as reflected by a 2×10^6-fold decrease in the affinity of the inhibitor for papain on truncation of N-terminal residues after the conserved Gly-9. Similar decreases in the affinity for proteinases were reported for cystatin C lacking the N-terminal residues preceding the conserved Gly-residue (Abrahamson et al. 1991; Björk et al., 1994). These studies also concluded that the role of the N-terminal segment of cystatin C and chicken cystatin (both family II cystatins) in the interaction differed between target proteinases. In the binding of the inhibitors to enzymes with accessible active sites, such as papain and actinidin, the N-terminal region thus primarily conferred stability to the final complex. However, the suggestion was made that the N-terminal region was required for the initial binding step of the proposed two-step binding of the cystatins to cathepsin B (Björk et al., 1994, 1996), whose active site cleft is partially covered by the occluding loop (Musil et al., 1991).

Several studies have approached how the N-terminal region of cystatins participates in interactions with target proteinases by genetically engineering substitutions in this region and monitoring changes in the affinity and kinetics of inhibition of the enzymes. Mutations of the evolutionarily conserved Gly to Glu or Trp drastically reduced the affinity of cystatin C for papain, ficin, actinidin and cathepsin B (Björk et al., 1995), and replacement of this residue also decreased the affinity of cystatin A for papain (Shibuya et al., 1995a). Moreover, Björk et al. (1995) observed the same differential kinetic behaviour for the interaction of the cystatin C mutants with the different enzymes as those reported for the N-terminally truncated forms of cystatin C and chicken cystatin (Abrahamson et al. 1991; Björk et al., 1994).

Mutation of residues N-terminal of the conserved Gly have primarily provided information regarding the specificity of the inhibitors for certain proteinases. A mutant of cystatin C, in which Val-10 was replaced by Arg, displayed only two-fold lower affinity for cathepsin B than the wild-type inhibitor, whereas it inhibited papain 12 000-fold more weakly than the wild-type (Lindahl et al., 1994). By investigating the effects of replacements of Leu-9 and/or Val-10 with Gly residues, Hall et al. (1995) concluded that most of the contributions

of the N-terminal region to the affinity of Trp-106 to Gly variants of cystatin C for cathepsins B, H and L were conferred by Val-10. However, the selectivity of these cystatin C variants for the enzymes studied resided in Leu-9. In subsequent work, the inhibition profile of cystatin C towards cathepsins B, H, L and S was altered by various replacements of these residues (Mason et al., 1998). Such changed inhibition profiles towards cathepsins B, H, L and S were also observed for hybrids of cystatin C with the N-terminal region of cystatin D, and vice versa (Hall et al., 1998). These results are consistent with the three-dimensional structure of the stefin B-papain complex (Stubbs et al., 1990), in which Ser-3 and Met-2 interact in a substrate-like manner with the unprimed S_2 and S_3 subsites, respectively, of the proteinase.

The effect of mutations in the highly conserved QVVAG sequence, i.e. the first binding loop, of cystatins for the affinity of the inhibitors for target proteinases has also been investigated, giving apparently contradicting results. Substitutions of Gln-46 to Lys and of Val-48 to Thr was reported not to detectably decrease the affinity of cystatin A for papain, cathepsin B or cathepsin L (Nikawa et al., 1989). This claim was in agreement with studies of cystatin B, in which replacement of Val-47 with one of several residues did not appreciably lower the affinity of the inhibitor for papain (Jerala et al., 1990). In dispute with these results, however, was a subsequent report by Machleidt et al. (1991), in which replacement of Val-48 in cystatin B by Asn was shown to result in a 240-fold drop of the affinity of the inhibitor for papain. Similar investigations of the first binding loop of chicken cystatin, in which Gln-53 was replaced by Glu or Asn, showed decreased inhibitory activity of the mutants for actinidin and cathepsin B but not for papain (Genenger et al., 1991). These results were partially revised later by the same group (Auerswald et al., 1992) with the use of kinetic measurements. In this work they reported 10- to 1000-fold decreased affinities for papain, actinidin and cathepsin B by several mutations in the first binding loop of chicken cystatin, e.g. the single amino acid replacements of Gln-53 to Glu and Asn and of Gly-57 to Ala, and the double exchanges of Arg-52 and Gln-53 to Leu and Glu, of Gln-53 and Ser-56 to Asn and Ala and of Leu-54 and Gly-57 to Met and Ala, respectively. In agreement with these results, variants of recombinant chicken cystatin lacking parts of the first binding loop bound more weakly than the wild-type inhibitor to papain, cathepsin B and cathepsin L (Auerswald et al., 1995). Interestingly, it was found that recombinant variants of chicken cystatin lacking parts of the first or second binding loops gave temporary inhibition of papain and cathepsin L, i.e. free enzyme and cleaved inhibitor were released from the enzyme-inhibitor complex after prolonged incubation (Machleidt et al., 1995). The variants of chicken cystatin were cleaved between the conserved Gly-9 and Ala-10 residues, indicating that disorders in the binding loops of cystatin affect the binding of the N-terminal region in such a way that it becomes accessible to an attack from the catalytically active Cys of the proteinase (Machleidt et al., 1995).

The importance of the second binding loop in the binding of cystatins to target proteinases has been studied for family II inhibitors, both by chemical modifications of the well conserved Trp residue and by engineered modifications of recombinant inhibitors. Early evidence for involvement of this region of the inhibitors in proteinase binding was presented by Nycander and Björk (1990). Chemical modification of the sole Trp, at position 104, of chicken cystatin with a tryptophan-specific reagent produced a 4×10^5-fold weaker inhibitor of papain than intact cystatin. Kinetically, this decrease in affinity was due essentially only to an increased dissociation rate constant. These results are in agreement with conclusions from a subsequent investigation in which the effects of deletion of the second binding loop of chicken cystatin were studied (Auerswald et al., 1995). In this work mutants lacking residues Pro-103–Leu-105 and Ile-102–Gln-107 displayed ~1000- and ~60-fold lower affinities for papain and cathepsin B, respectively, although only ~10-fold lower affinities for cathepsin L (Auerswald et al., 1995). The decreases in affinity arose predominantly from increased dissociation rate constants, suggesting an anchoring role for this region in the binding of chicken cystatin to target proteinases.

Studies conducted on recombinant cystatin C with mutations of the conserved Trp of this inhibitor, Trp-106, also showed the importance of the second binding loop for high affinity binding of the inhibitor to target proteinases. Hall et al. (1995) reported more than 10-fold decreased affinities of a Trp-106 to Gly variant of cystatin C for cathepsins B, H, L and S. In later work, it was shown that mutations of Trp-106 to Gly resulted in 300–900-fold lowered affinities of the mutant for papain, actinidin and cathepsins B, H and L (Björk et al., 1996). However, smaller decreases in affinities for the enzymes were observed for a Trp-106 to Phe mutant, indicating that the phenyl side chain could to some extent compensate for the absence of the indole side chain of Trp-106, and providing evidence for the hydrophobic nature of the inhibitor-enzyme interactions (Stubbs et al., 1990). Characterization of the kinetics of binding of these mutants to target enzymes showed that the lowered affinities were essentially due only to increased dissociation rate constants.

General conclusions regarding binding mechanism from previous investigations

Family II cystatins

This section will begin by a discussion of the mechanism of binding of cystatins of family II to target proteinases, due to the larger number of investigations dealing with cystatins of this family. On the basis of structural data, and of chemical modifications and engineered mutations of the three

binding regions of family II inhibitors, the importance of these regions for proteinase binding is quite well established. The structure of the inhibitory wedge-shaped edge formed by these three regions is presumably complementary to the active site cleft of papain-like proteinases with accessible clefts, such as papain. This implicates that a tight fit between inhibitor and enzyme is achieved without major conformational changes. Kinetic data strongly support this assumption by demonstrating that the association reaction is consistent with the formation of the enzyme-inhibitor complex in a single step, and that the bimolecular association rate constant for this process is as fast as would be expected for a diffusion-controlled reaction. However, no structure of the same free and complex- bound inhibitor is available to indicate what changes, if any, occur. The proposed model for the interaction between the enzyme and the inhibitor is that the two hairpin loops bind to the active site cleft of the enzyme, forming many primarily hydrophobic and rather unspecific interactions with the primed substrate-binding subsites of the enzyme. The N-terminal region plays a major role in the binding by anchoring the inhibitor to the unprimed subsites of the enzymes, and by conferring specificity for certain enzymes to the cystatins. The anchoring is achieved primarily by interaction of the residue immediately preceding the conserved N-terminal Gly residue with the S_2 subsite of the enzyme. The three-dimensional solution structures of family II cystatins reveal that the configuration of the two hairpin loops is quite rigid, in contrast to the region preceding the conserved N-terminal Gly residue, which is very flexible. These findings, together with the demonstration that the N-terminal region is dispensable for the rapid association of cystatins of family II with enzymes with accessible active site clefts, indicate that this region may bind in a second step, although too fast to have been detected experimentally.

A somewhat different mechanism appears to govern the binding of the inhibitors to cysteine proteinases of the papain family that do not have accessible active site clefts, of which cathepsin B is the only enzyme that has been extensively investigated. In order for the inhibitor to form a tight complex with cathepsin B, the occluding loop that partially covers the active site cleft of the enzyme has to be displaced. This is reflected by the study of the binding kinetics of cystatin C to cathepsin B, which indicates a two-step binding mechanism. In this binding process, the N-terminal region of cystatin is needed for the rapid association of the interacting proteins, indicated by lowered association rate constants with cathepsin B of cystatin variants lacking their N-terminal regions. The suggested mechanism is thus that an initial, weak complex is formed between the N-terminal region of the cystatin and cathepsin B, which enables the rest of the inhibitory wedge to displace the occluding loop of proteinase, and to gain access to the active site cleft of the enzyme, thereby forming a more stable enzyme-inhibitor complex.

Family I cystatins

The mechanism by which cystatins of family I interact with cysteine proteinases of the papain family has been less investigated than that of family II cystatins. This relative lack of data is especially evident in the case of the role of the three binding regions of the inhibitor. Contradictory results, which probably reflect difficulties in quantifying tight affinities, disputing whether the N-terminal region of cystatins of family I contributes or not to the binding of the inhibitors to target proteinases make it difficult to assess the role of this region. The limited number of investigations dealing with the contribution of the first binding loop of family I cystatins to the binding have also given somewhat contradictory results, although they mostly indicate that this region is of similar importance for the inhibitory activity as it is for cystatins of family II. Moreover, no studies are available for the role of the second binding loop. However, all the three regions are implicated in interactions with the enzyme in the crystal structure of cystatin B in complex with papain. Furthermore, this crystal structure and the solution structure of cystatin A reveal a similar overall structure between cystatins of families I and II, which indicates that the binding mechanism of cystatins of the two families to target proteinases should be very similar. The few investigations done on the binding kinetics of family I cystatins to papain-like cysteine proteinases with accessible active site clefts confirm the latter assumption, and support the model of these cystatins binding in a one-step mechanism. Despite these limited and somewhat contradictory results, it is reasonable to assume that the mechanism of interaction between cysteine proteinases of the papain family and cystatins of families I or II should essentially be the same, i.e. that proposed above for family II cystatins.

Present investigation

This thesis work was conducted with the purpose of elucidating the mechanism that governs the interactions between cysteine proteinases of the papain family and cystatin A, a representative of family I cystatins. In particular, the role of the N-terminal region of the inhibitor in these interactions was investigated.

Characterization of recombinant cystatin A (Paper I)

When this work was started, endogenous cystatin A had been isolated from several sources (Brzin et al., 1983; Green et al., 1984), and recombinant inhibitors had been expressed in bacterial systems (Fong et al., 1989; Nikawa et al., 1989; Kaji et al., 1990; Jerala et al., 1994). However, no extensive characterization of the interactions of cystatin A with cysteine proteinases had been reported. An expression system generating substantial quantities of fully active inhibitor was developed in this study, which enabled analyses by spectroscopic, kinetic and equilibrium methods of the interactions between the inhibitor and several target proteinases.

Analyses of the near-UV absorption, circular dichroism and fluorescence changes accompanying the binding of cystatin A to papain indicated that these changes arose predominantly from perturbations of the environment of tryptophan residues. The lack of tryptophans in cystatin A implicated residues in papain as responsible for these spectroscopic changes, Trp-177 and/or Trp-69 of the enzyme being the most likely candidates, due to their close vicinity to the active site cleft. The spectral changes for the binding of cystatin A to papain were appreciably different from those observed for the interactions of cystatin C or chicken cystatin with the enzyme (Lindahl et al., 1988). This finding is consistent with the absence in cystatin A and presence in chicken cystatin and cystatin C of a tryptophan residue in the second binding loop that has been proposed to contribute substantially to the spectral changes seen on binding of the latter inhibitors to papain (Lindahl et al., 1988, 1992a).

The affinities of cystatin A for papain and cathepsin L were found to be 50–100-fold higher and the affinity for cathepsin B five-fold higher than in earlier reports (Green et al., 1984; Fong et al., 1989; Nikawa et al., 1989; Abrahamson et al., 1986; Kaji et al., 1989). These discrepancies most certainly are due to difficulties in accurately measuring the tight interactions of cystatin A with these target proteinases. The affinities of the recombinant cystatin A for the physiological target enzymes cathepsins B and L, which apparently had been appreciably underestimated in previous work, were in better agreement with the values reported for bovine cystatin A (Turk et al., 1995b). In contrast with these findings, the affinities of recombinant cystatin A for cathepsins C and H were comparable with those determined in previous studies (Barrett et

al., 1986; Abrahamson et al., 1991; Hall et al., 1993; Nicklin and Barrett, 1984).

Of the enzymes investigated, papain and cathepsin L were the most strongly inhibited by cystatin A, with K_i values of 0.2 and 20 pM, respectively. The affinity of cystatin A for papain is thus about one order of magnitude lower than that of chicken cystatin or cystatin C, both belonging to family II (Björk et al., 1989; Lindahl et al., 1992a), but considerably higher than that reported for cystatin B (Jerala et al., 1994), another family I cystatin. Cathepsin L is inhibited with similar affinities by cystatin A and cystatin B (Jerala et al., 1994), whereas cystatin C and chicken cystatin appear to have higher affinity for this enzyme (Anastasi et al., 1983; Barrett et al., 1984; Green et al., 1984). The association rate constants, k_{ass}, for the binding of cystatin A to papain and cathepsin L were in the range of $3\text{-}5 \times 10^6$ $M^{-1} \cdot s^{-1}$, and comparable with k_{ass} values determined for the interactions of several cystatins with these enzymes. Such k_{ass} values are close to those expected for a diffusion-controlled reaction. Moreover, analyses of the kinetics of binding of cystatin A to papain at high inhibitor concentrations revealed no deviation from a linear dependence of the observed pseudo-first order rate constant, k_{obs}, on inhibitor concentration. The data are thus consistent with the formation of the papain-cystatin A complex in one step, without major conformational changes of either protein, as proposed previously for other cystatins (Bode et al., 1988; Björk et al., 1989; Lindahl et al., 1992a).

K_i values for the inhibition of the other enzymes studied in this work, i.e. cathepsins B, C and H and actinidin, by cystatin A were in the range of 1–40 nM, thus reflecting appreciably weaker affinities than those for papain and cathepsin L. The K_i of cystatin A for cathepsin B is comparable with the values of chicken cystatin and cystatin C for this enzyme, but significantly higher than that of cystatin B (Green et al., 1984; Björk et al., 1994; Abrahamson et al., 1991; Jerala et al., 1994; Hall et al., 1993). However, the lower affinities of cystatin A and the two family II inhibitors for cathepsin B have different kinetic backgrounds. The reduced affinity of cystatin A for cathepsin B compared with that for papain, is thus due to a reduced k_{ass}, whereas the decreased affinity of the family II cystatins for this enzyme instead reflects primarily increased k_{diss}. Cystatin A binds to actinidin with an affinity comparable with those of chicken cystatin and cystatin C, but the latter cystatins appear to be 10–100 fold better than cystatin A in inhibiting cathepsins C and H (Nicklin and Barrett, 1984; Björk et al., 1990; Lindahl et al., 1992a; Barrett 1986; Abrahamson et al., 1991; Hall et al., 1993). Also cystatin B binds with 10–50 fold higher affinity to cathepsins C and H than cystatin A (Green et al., 1984; Jerala et al., 1994).

Cystatin A was found to bind to inactivated papains, although with decreasing affinities as the size of the inactivating group increased. This decrease was entirely due to increased k_{diss}, as reported for similar studies with family II

cystatins (Björk and Ylinenjärvi, 1989; Lindahl et al., 1992a). However, much larger effects were obtained for cystatin A, indicating that there is less space for an inactivating group around the active site Cys-25 of papain in the papain-cystatin A complex than in complexes of the enzyme with chicken cystatin or cystatin C.

A truncated variant of cystatin A, lacking the first six residues, displayed a substantially (i.e. ~7 × 10^6-fold) lower affinity for papain than the wild type inhibitor. This observation was in accordance with the lower inhibitory capacity observed for truncated variants of rat cystatin α (Takeda et al., 1986), and suggested a similar importance for the N-terminal region of cystatin A in proteinase binding as for the corresponding region in chicken cystatin and cystatin C (Lindahl et al., 1992b; Björk et al., 1994, 1995; Abrahamson et al., 1991; Machleidt et al., 1989; Hall et al., 1993). However, the affinity of the interaction was not accurately determined, and the somewhat altered CD spectrum of the variant was also consistent with an altered conformation of the inhibitor being responsible for the reduced affinity.

The role of Gly-4 of cystatin A in proteinase binding (Paper II)

At the outset of this investigation, the importance of the evolutionarily conserved Gly residue in the N-terminal region of cystatins for the interactions with target proteinases had been extensively investigated for family II cystatins (Hall et al., 1998; Abrahamson et al., 1991; Machleidt et al., 1989; Lindahl et al., 1992b; Björk et al., 1994, 1996). However, only one such study of family I cystatins had been reported (Shibuya et al., 1995a). This investigation was limited to exploration of how the affinity of Gly-4 mutants of cystatin A for the model enzyme, papain, decreased depending on the size of the side chain of the replacing amino acid. Moreover, the reported affinities were uncertain, due to the methodology used. Consequently, very little was known about the magnitude of the effects and the underlying mechanism responsible for the lowered affinities. Also, no data were available for interactions of cystatin A Gly-4 mutants with physiologically relevant proteinases, such as cathepsins B or L. In order to assess the importance of the evolutionarily conserved Gly-4 of cystatin A, several Gly-4 mutants were expressed in *E. coli*, and their ability to inhibit papain and cathepsins B and L, as well as their kinetics of binding to these enzymes, were characterized in detail. In addition, the structural changes resulting from these mutations were explored by analyses of the CD spectra and by NMR studies of the mutants.

In general, the affinities of the cystatin A Gly-4 variants for the enzymes studied in this work, i.e. papain and cathepsins B and L, decreased with the size of the side chain of the substituting residue, as reported earlier for papain (Shibuya et al., 1995a). However, the magnitudes of these lowered affinities were in strong contrast with those obtained in the previous study (Shibuya et

al., 1995a). In the present work, even the substitution of Gly-4 to Ala or Ser resulted in 900- and 7000-fold decreases, respectively, of the affinity of the inhibitor for papain, from an affinity of the wild-type inhibitor of ~0.2 pM. In contrast, Shibuya et al. (1995a) reported only 10- to 20-fold reduced affinities for these mutations. Moreover, the K_i values for papain of cystatin A variants in which Gly-4 was replaced by Arg, Glu or Trp increased to ~1 μM. These substitutions thus reduced the affinity by more than 3×10^6-fold, whereas the maximal decrease reported in the previous work was 240-fold (Shibuya et al., 1995a). These conflicting results between the two investigations are most evident for cystatin A variants displaying high affinities for papain and are likely to reflect the inappropriate methodology used previously for determinations of such high affinities (Shibuya et al., 1995a).

The affinity for cathepsin L of the Gly-4 mutants of cystatin A also decreased with increasing size of the side chain of the substituent. The K_i values of the mutants for this enzyme were similar to those measured for papain, with one exception. The affinity of the Gly-4 to Trp mutant for cathepsin L was thus 50-fold higher than that for papain, presumably reflecting specific interactions of the indole group in position 4 of the inhibitor with cathepsin L. In general, the decreased affinities of the Gly-4 cystatin A mutants for cathepsin L, relative to that of the wild-type, were less pronounced than for papain. This could partially be due to the somewhat uncertain K_i value of wild-type cystatin A for cathepsin L that was used for comparison. This K_i value (~20 pM) was determined previously by equilibrium measurements, as lack of material precluded more accurate separate determinations of the association and dissociation rate constants in the way done for papain (Pol et al., 1995). Another explanation of these results is that cathepsin L better than papain tolerates other residues than Gly in position 4 of cystatin A.

In the case of cathepsin B, even more drastic effects on the affinity of the Gly-4 mutants for the enzyme were observed. Even the smallest change, substitution by Ala, resulted in a 7000-fold decreased affinity for cathepsin B, compared with that of wild-type cystatin A. The Gly-4 to Ser variant showed another 4-fold lowered affinity, and replacement of Gly-4 by Arg, Glu or Trp resulted in very weak inhibitors of cathepsin B, for which K_i was indeterminable, although estimated to be > 100 μM.

As the decreased binding affinities of the Gly-4 cystatin A mutants for target proteinases may have arisen from conformational changes of the inhibitors induced by the mutations, the possibility of such changes were analysed by far-UV CD and NMR. No conformational changes were evident from the experimentally indistinguishable CD spectra of the cystatin A variants and that of the wild-type inhibitor. Moreover, NMR analyses of the Gly-4 to Ala and Gly-4 to Trp mutants showed only minimal changes of chemical shifts for most of the residues. For both mutants, the largest changes in chemical shift were for residues in immediate vicinity of the mutated residue, such as Gly-5

and Leu-6, but also for Val-47 in the first binding loop. Although detectable, the magnitudes were very small (< 0.2 ppm), and provided no evidence that the structure of cystatin A was affected by the mutations . These findings thus indicate that the decreased binding affinities of the mutants for proteinases are not due to altered conformations of the inhibitors, but arise from perturbations of the interactions between the enzyme and the mutated side chain in position 4.

Data on the changes of binding kinetics responsible for the weaker interaction of the mutants with target proteinases could be obtained for all three enzymes only for the cystatin A variants in which Gly-4 was replaced by Ala or Ser. The results for these mutants were clear and consistent with similar studies of family II cystatins (Abrahamson et al., 1991; Lindahl et al., 1992b; Björk et al., 1994, 1996). The reduced affinities were due only to increased dissociation rate constants in the case of papain and cathepsin L, whereas both decreased association and increased dissociation rate constants contributed to the reduction in affinity for cathepsin B. These findings indicate that the region around the Gly residue is only involved in the stabilization of the complex between the inhibitor and papain or cathepsin L. In contrast, this region is also needed for the association step of the interaction of the inhibitor with cathepsin B. As discussed further below (Paper III), this behaviour presumably reflects the requirement for an initial binding of the N-terminal region of cystatin A to cathepsin B, in order to facilitate the subsequent displacement of the occluding loop of the enzyme. Such displacement has to occur before the rest of the proteinase binding regions of the inhibitor can gain access to the partially obstructed active site cleft of cathepsin B and form a tight complex (Musil et al., 1991; Björk et al., 1994, 1996; Nycander et al., 1998).

Analyses of the binding kinetics of the Gly-4 to Arg and Gly-4 to Glu mutants to papain showed an hyperbolic dependence of the observed pseudo-first order rate constant, k_{obs}, on inhibitor concentration. This behaviour is that expected for the formation of the enzyme-inhibitor complex via a two-step mechanism, in which the free proteins (E and I) establish a rapid equilibrium with an initial, weak complex (EI) before the more stable, final complex (EI*) is formed by a conformational change, as shown in Scheme 2.

$$\mathrm{E} + \mathrm{I} \underset{}{\overset{K_1}{\rightleftarrows}} \mathrm{EI} \underset{k_{-2}}{\overset{k_{+2}}{\rightleftarrows}} \mathrm{EI}^* \qquad \text{(Scheme 2)}$$

In this scheme, K_1 is the dissociation equilibrium constant for the first step, and k_{+2} and k_{-2} are the rate constants for the interconversion between the two forms of the complex. Data for the dependence of k_{obs} on the concentration

of both cystatin A variants were satisfactorily fitted to the equation for this mechanism (Bieth, 1995). This observed two-step binding of cystatin A variants with charged substituents in position 4 to papain indicates that these side chains cannot be accommodated unless conformational changes occur in the complex. Such internal rearrangements of complexes between target proteinases and other cystatin A Gly-4 variants may also occur, although not detectable by the methodology used.

Binding of N-terminal truncation variants of cystatin A to proteinases (Paper III)

The two earlier studies (Papers I and II), as well as investigations by other groups (Takeda et al., 1985; Shibuya et al., 1995a, b), indicated that the N-terminal region of cystatin A plays a substantial role in the tight binding of the inhibitor to target proteinases. However, data regarding the contribution and importance of individual residues preceding the evolutionarily conserved Gly-4 residue to the interaction with papain and physiologically relevant cathepsins were lacking, or were uncertain.

In this work, mutants of cystatin A were produced, in which the three N-terminal residues were sequentially deleted. Such truncation resulted in lowered affinities of the variants for the enzymes studied, i.e. papain and cathepsins B and L. The solution structure of cystatin A (Martin et al., 1995) shows a high flexibility of the N-terminal region, and a lack of interactions of residues preceding Leu-6 with the rest of the protein. These observations, together with analyses of the CD spectra of the truncation variants in this work, in which no alterations were detectable, indicate that the truncation of the residues from the N-terminus of cystatin A is not accompanied by any conformational changes of the inhibitor. The decreases in binding affinity for proteinases of the variants lacking N-terminal residues are thus likely to entirely reflect the absence of contributions of these residues to the binding.

The deletion of Met-1 had no detectable effects on the affinities of cystatin A for papain and cathepsin L, and only resulted in a 4-fold decrease in the affinity for cathepsin B, indicating that this residue is of minor importance for the inhibitory capacity of cystatin A. This finding is in agreement with Met-1 not been involved in interactions with papain in the crystal structure of the complex between papain and cystatin B (Stubbs et al., 1990), a homologous family I cystatin. Further truncation, of Met-1 and Ile-2, had more serious consequences for the inhibitory ability. Compared with the wild-type inhibitor, this variant had 900- and 200-fold lower binding affinity for papain and cathepsin B, respectively, whereas the decrease for cathepsin L could only be estimated to be larger than 3-fold, due to the uncertain value of K_i obtained for wild-type cystatin A. The affinities for cathepsin L, papain and cathepsin B of the mutant lacking the three residues preceding the conserved Gly-4 were

reduced another 20 000-, 2000- and 400-fold, respectively, indicating the predominant contribution of Pro-3 to the high affinity binding of cystatin A to the three proteinases. The reduced affinities of this truncation variant of cystatin A for papain and cathepsin B indicate that the N-terminal region contributes about 40% of the total free energy of binding of the intact inhibitor to the two enzymes, a contribution similar to that reported for chicken cystatin (Lindahl et al., 1992b).

These results show conclusively that the N-terminal region of cystatin A is indispensable for the high affinity binding of the inhibitor to target proteinases, being as important as the corresponding region in family II cystatins is for such binding (Hall et al., 1998; Abrahamson et al., 1991; Machleidt et al., 1989; Lindahl et al., 1992b; Björk et al., 1994; Björk et al., 1996). They also show that most of the contribution of the N-terminal region to the affinity for proteinases is conferred by the two residues adjacent to the evolutionarily conserved Gly-4 residue, i.e. Pro-3 and Ile-2, primarily the former. A previous study had reported that truncation of the two first residues of a Met-65 to Leu cystatin A variant only had a minimal effect, ~3-fold, on the affinity of this variant for papain (Shibuya et al., 1995b), whereas the same deletion gave a ~900-fold reduced affinity in the present work. Moreover, in the previous study (Shibuya et al., 1995b), it was claimed that the ability of cystatin A to inhibit papain only was compromised for truncations beyond Pro-3 and that deletion of the first four residues, including Gly-4, resulted in a ~200-fold weaker inhibitor of papain. In contrast, deletion of the residues preceding Gly-4 in this work gave ~2×10^6-fold decreased affinity for papain compared with the wild-type inhibitor. These apparent discrepancies are most likely due to difficulties in measuring the very low inhibition constants of some of the cystatin A variants for papain by equilibrium methods. Therefore, many of the conclusions of the previous work (Shibuya et al., 1995b) concerning the role of the N-terminal region of cystatin in proteinase binding that are based on such uncertain K_i values must be considered questionable.

Analyses of the kinetics of inhibition of the three enzymes studied in this work by the N-terminally truncated cystatin A variants showed that the truncations did not affect the rate by which the inhibitors bound to papain or cathepsin L. The reduced affinities of the truncated variants for these enzymes was instead only caused by an increased k_{diss}, i.e. less stable complexes were formed. In contrast, both a decreased k_{ass} and an increased k_{diss} were responsible for the reduced affinities of the cystatin A variants for cathepsin B. Similar results has been obtained in studies of the kinetics of inhibition of these enzymes by N-terminally truncated variants of the family II cystatins, chicken cystatin and cystatin C (Abrahamson et al., 1991; Lindahl et al., 1992b; Björk et al., 1994; Björk et al., 1996).

On basis of these and other results, it can be concluded that the N-terminal region of cystatin A binds to papain and cathepsin L in a substrate-like

manner, anchoring the inhibitor to the enzyme by interactions of Ile-2 and Pro-3 with the S_3 and S_2 substrate binding subsites of the enzyme, respectively (Stubbs et al., 1990; Lindahl et al., 1994; Hall et al., 1995; Mason et al., 1998; Hall et al., 1998), but not affecting the rate of the reaction. Given the flexibility of the N-terminal region of cystatin A, and other cystatins, in solution (Dieckmann et al., 1993; Martin et al., 1995; Ekiel et al., 1997), is possible that this binding occurs subsequent to that of the more rigid first and second binding loops of the inhibitor. However, no indications of such a two-step binding mechanism have been detected by rapid-kinetic analyses of the interaction of several cystatins with papain and other proteinases with accessible active site clefts (Björk et al., 1989; Björk et al., 1990; Lindahl et al., 1992a; Pol et al., 1995; Turk et al., 1994; Turk et al., 1995b).

In contrast, cystatin A probably binds by a different mechanism to cathepsin B, in which the active-site cleft is partially obstructed by the occluding loop (Musil et al., 1991). The association of cystatin C with cathepsin B has been shown to occur via a two-step mechanism (Nycander et al., 1998), presumably reflecting the need of the inhibitor to displace this occluding loop before gaining access to the active site of the enzyme. The decreased k_{ass} for cathepsin B of cystatin A variants lacking the N-terminal region, as well as of Gly-4 mutants of the inhibitor (Paper II), indicates that the N-terminal region of cystatin A is involved in the formation of the initial complex of such a two-step reaction. The displacement of the occluding loop is facilitated by the formation of this initial complex, allowing the subsequent interaction of the two other binding regions of the inhibitor with the active site cleft of the proteinase. A similar role for the N-terminal region in interactions of family II cystatins with cathepsin B has been proposed from studies with N-terminally modified such inhibitors (Björk et al., 1994; Björk et al., 1995, Björk et al., 1996).

Two-step proteinase binding of N-terminally labelled cystatin A (Paper IV)

Studies of the kinetics of binding of cystatin A to papain have shown that the association of the two proteins is very rapid, with a k_{ass} of ~3×10^6 $M^{-1} \cdot s^{-1}$, and experimentally consistent with a simple bimolecular reaction involving no kinetically detectable conformational changes (Pol et al., 1995). Such a one-step binding mechanism of cystatins to cysteine proteinases has also been proposed for the interactions of chicken cystatin and cystatin C with papain and similar enzymes (Björk et al., 1989; Björk and Ylinenjärvi, 1990; Lindahl et al., 1992a) and is supported by the highly complementary shape of the interacting regions of chicken cystatin to the active-site cleft of papain (Bode et al., 1988). However, the high flexibility in solution of the N-terminal region of cystatins (Dieckmann et al., 1993; Martin et al., 1995; Ekiel et al., 1997),

together with the unaffected k_{ass} for the binding of N-terminally truncated cystatins to papain (Lindahl et al., 1992b; Björk et al 1994, 1996; Paper III), make it plausible that the N-terminal region may bind to the enzyme in a second step, after the binding of the two, more rigid, loops of the inhibitor.

To allow the kinetics of binding to papain of the N-terminal region of cystatin A to be differentiated from those of the other regions of the inhibitor, cystatin A was expressed with an extra N-terminal Cys residue, which was labelled with a fluorescent group. The interactions of the labelled inhibitor with papain were studied by stopped-flow methods, either by monitoring the decrease of enzyme activity or by following the increase of the fluorescence of the label that resulted from interactions with the enzyme. In the former analyses, the dependence of k_{obs} on the concentration of the inhibitor, present in excess, was linear, giving a k_{ass} of ~2×10^7 $M^{-1} \cdot s^{-1}$ and a k_{diss} indistinguishable from zero. In contrast with this linear behaviour, k_{obs}, monitored by the interaction of the N-terminal label with papain, showed the same hyperbolic dependence on reactant concentration, approaching a limiting rate constant at high reactant concentrations, regardeless of whether the enzyme or the inhibitor was in excess. These different behaviours of k_{obs} depending on which monitoring method was used are compatible with a two-step binding mechanism, in which the proteinase (P) and the rigid regions of the binding edge of the inhibitor (I) form an initial complex (PI), in which the proteinase is prevented from cleaving substrate. The second step of this mechanism involves the binding of the labelled N-terminal region of the inhibitor to the enzyme (PI*), which stabilizes the complex, alters the fluorescence of the label and accounts for the saturation kinetics observed in experiments monitored by this property:

$$P + I \underset{k_{-1}}{\overset{k_{+1}}{\rightleftarrows}} PI \underset{k_{-2}}{\overset{k_{+2}}{\rightleftarrows}} PI^*$$

From the hyperbolic dependence of k_{obs} on reactant concentration, values for $K_{0.5}$ and k_{lim} of ~10 μM and ~230 s^{-1}, respectively, were derived. $K_{0.5}$ is the concentration of the reactant in excess at which half of the limiting rate constant is reached and k_{lim} is this rate constant. k_{lim} must reflect the rate constant,k_{+2}, for the convertion of the intermediate to the final complex, therefore implicating that k_{+2} is ~230 s^{-1}. From the linear dependence of k_{obs} obtained from the decrease of enzyme activity on inhibitor concentration which reflects the first step of the reaction, and the intercept on the ordinate of this plot, if follows that $k_{-1} << k_{+2}$. An estimated value for k_{-1} is ~1 s^{-1}, which is the k_{diss} for a complex of papain with cystatin A lacking the N-terminal region (Paper III). Under these conditions, the initial slope of the hyperbolic dependence, k_{ass}, corresponds to k_{+1}, and k_{+1} is thus ~2×10^7 $M^{-1} \cdot s^{-1}$. This

value is identical to that obtained by monitoring the first step of the reaction by measurements of the rate of disappearance of the free enzyme. Finally, k_{-2}, can be estimated from the overall K_d, calculated from k_{ass} and a separately determined k_{diss} (~5×10^{-7} s^{-1}), together with the values of k_{+1}, k_{-1} and k_{+2} and discused above. This gives a value of ~10^{-4} s^{-1} for k_{-2}, the rate constant fordissociation of the N-terminal region of the labelled inhibitor from the proteinase.

The almost identical k_{diss} values obtained for the interactions of wild-type cystatin A (Pol et al., 1995) and the labelled variant with papain indicate that the label is not involved in the binding but acts merely as a reporter of environmental changes around the N-terminus of the inhibitor. This conclusion is further supported by the structure of the cystatin B-papain complex, in which Met-1 of the inhibitor is not involved in interactions with the enzyme (Stubbs et al., 1990), and by the unaffected affinity for papain of a cystatin A variant lacking Met-1 (Paper III). These observations provide strong evidence that the second step of the two-step mechanism of binding of the labelled cystatin A to papain reflects the binding of the N-terminal region and not simply a binding or rearrangement of the label.

Previous fast-kinetic studies of interactions of cystatin A, or other cystatins, with papain have not indicated that a two-step mechanism is involved in the reactions investigated. One explanation may be that the changes in tryptophan fluorescence used to monitor the reactions in the previous work only report the first step of a two-step binding of the inhibitor to papain, in analogy with the kinetic studies in the presence of substrate in this work. This possibility is supported by the structure of the cystatin B-papain complex (Stubbs et al., 1990), in which the Trp-177 residue in papain, proposed to account for the majority of the fluorescence changes during complex formation (Lindahl et al., 1988; Pol et al., 1995), is located in the vicinity of the two loops, and further away from the N-terminal region of the inhibitor. Another possibility is that the unlabelled N-terminal region of cystatin A binds to the enzyme with a considerably faster rate constant than the labelled region. A six-fold increased k_{+2}, from ~230 s^{-1} for the labelled inhibitor to 1500 s^{-1} for the unlabelled inhibitor, together with the lower k_{ass} of ~3×10^6 $M^{-1} \cdot s^{-1}$ for the latter (Pol et al., 1995), would give a $K_{0.5}$ of 500 μM. In this case, even if the binding of the N-terminal region of the unlabelled inhibitor to papain results in changes of tryptophan fluorescence, the high $K_{0.5}$ would give an apparently linear dependence of k_{obs}, up to 100 μM, which was the highest inhibitor concentration that could be studied in the previous work (Pol et al., 1995).

Conclusions

The aim of this work has been to elucidate of the detailed mechanism by which cystatins bind to and inhibit their target enzymes, the papain-like cysteine proteinases. In particular, the interactions of recombinant wild-type- and mutated cystatin A variants with such proteinases were investigated by spectroscopic, kinetic and equilibrium methods.

The mammalian cysteine proteinase inhibitor, cystatin A, was expressed in bacteria and purified to give fully active protein. The binding of cystatin A to several proteinases of the papain family was quantified, and it was found that the affinities for papain and cathepsins B and L had been largely underestimated in previous work. Papain and cathepsin L were inhibited rapidly and strongly, with k_{ass} values of ~3–5 × 10^6 $M^{-1} \cdot s^{-1}$ and K_i values of ~0.2–20 pM. These values of k_{ass} approach those expected for a diffusion-controlled reaction. This observation, together with the linear concentration dependence of k_{obs} for the binding of cystatin A to papain at high inhibitor concentrations, is consistent with the association of the two proteins occurring in one step, without detectable conformational changes. Cystatin A was a less potent inhibitor of actinidin and cathepsins B, C and H, having K_i values for these enzymes in the range of 1–40 nM. These lower affinities were due primarily to an increased k_{diss}, except for cathepsin B, for which k_{ass} was ~100-fold lower than the values for the other enzymes.

The evolutionarily conserved Gly-4 residue of cystatin A was mutated to several other residues. The affinities of the inhibitor for papain and cathepsins B and L were substantially decreased even by the Gly-4 to Ala mutation. Moreover, replacement of Gly-4 by a charged residue, such as Glu or Arg, resulted in an affinity loss of more than five orders of magnitude for the three enzymes. The decreased affinities for papain and cathepsin L were due to an increased k_{diss}, whereas a combination of decreased k_{ass} and increased k_{diss} values were responsible for the lower affinities of the mutants for cathepsin B. This behaviour presumably reflects different roles of the N-terminal region of cystatin A in the binding to different enzymes. This region thus apparently serves mainly to anchor the inhibitor to proteinases with easily accessible active-site clefts, such as papain and cathepsin L, thereby reducing k_{diss}. In contrast, the N-terminal region of the inhibitor most likely is needed for the displacement of the occluding loop of cathepsin B and thereby increases the k_{ass} for the reaction with this enzyme. No evidence for structural changes of the mutants being responsible for the lowered binding affinities were detected by CD or NMR analyses.

Sequential deletion of residues from the N-terminal region of cystatin A revealed the importance and contribution of individual residues of this region

in the interaction of the inhibitor with papain and cathepsins B and L. The deletion of Met-1 minimally affected the inhibitory ability. Further deletion of Ile-2 and Pro-3 showed that these residues together conferred most of the affinity of binding of the N-terminal region to the enzymes and that Pro-3 was responsible for the major part of this affinity. The Ile-2 and Pro-3 residues presumably interact in a substrate-like manner with the S_3 and S_2 subsites of the enzymes, respectively. Deletion of all three residues before the evolutionarily conserved Gly-4 in a Δ MIP variant resulted in affinity losses of around five orders of magnitude for all three enzymes. This decreased affinity indicates that the N-terminal region provides about 40% of the free energy of binding of the inhibitor to the enzymes. The lower affinities of the truncated cystatin A variants for papain and cathepsin L are only the result of higher k_{diss} values, indicating, as for the Gly-4 mutations that the N-terminal region primarily has a role in anchoring the inhibitor to these enzymes. In contrast, the lowered k_{ass} values of the variants for cathepsin B similarly indicate that this region is needed for the formation of the initial complex between cystatin A and cathepsin B in the two-step binding mechanism of cystatins to this enzyme. As in the case of the Gly-4 mutants, no detectable changes in the CD spectra accompanied the truncations, providing evidence for conformational changes of the inhibitors not being responsible for the observed decreases in affinity.

Cystatin A was expressed with an extra N-terminal cysteine, which was labelled with a fluorescent group, AANS. AANS-cystatin A had eight-fold higher affinity for papain than the wild-type inhibitor. The k_{diss} of the labelled variant was indistinguishable from that of the wild-type, an increased k_{ass} thus being responsible for the higher affinity. k_{obs}, monitored by the loss of enzyme activity in the presence of substrate, increased linearly with inhibitor concentration and gave a k_{ass} of $\sim 2 \times 10^7\ M^{-1} \cdot s^{-1}$. In contrast, analyses of the interaction of AANS-cystatin A with papain, monitored by changes in the fluorescence of the label, showed that k_{obs}varied hyperbolically with reactant concentration. This behaviour indicates that the labelled N-terminal region of cystatin A binds to papain in the second step of a two-step reaction, after an initial binding of the rest of the inhibitory wedge to the enzyme. The k_{ass} derived for this two-step mechanism was identical with that obtained from measurements of the rate of inhibition of papain by the labelled inhibitor, i.e. $\sim 2 \times 10^7\ M^{-1} \cdot s^{-1}$. The limiting rate constant for the formation of the final complex, k_{+2}, was $\sim 230\ s^{-1}$, and the reactant concentration at which half of this limiting rate was reached was ~10 μM. It is likely that also the unlabelled N-terminal region of cystatin A binds to papain in a second step, although such binding could not be detected in previous kinetic studies monitored by changes of tryptophan fluorescence.

Acknowledgements

This investigation was carried out at the department of Veterinary Medical Chemistry, The Biomedical Center, Faculty of Veterinary Medicine, Swedish University of Agricultural Sciences, Uppsala.

I wish to thank:

My supervisor, Prof. Ingemar Björk, for his sense of responsibility and commitment in the guidance of his students, in this case myself. It is almost unnecessary to mention his outstanding scientific knowledge, but his generosity in sharing this knowledge in a pedagogic way deserves to be emphasized. I am very happy and proud to have had him as my tutor.

My closest colleagues, Mrs. Alona Pavlova, Dr. Maria Nycander and Mrs. Ewa Pol, for fruitful collaboration and for rewarding conversations regarding both scientific and more existential subjects.

All other co-authors, for fruitful collaboration.

Other members of the "IB-team", Dr. Véronique Arocas, Miss Karin Ylinenjärvi, Mrs. Kerstin Nordling, Mrs. Elke Raub-Segall and Mrs. Ingrid Brieditis, for creating a friendly athmosphere and for showing patience with my, sometimes, lack of good laboratory practice. Dr. Peter Lindahl and Dr. Charlotte Hallén-Sandgren , for helping a fresh and unsecure student. Dr. Thomas Prasthofer, for being so patient and encouraging while introducing me into the world of molecular biology.

Dr. Boris Turk, for being a friend and a source of inspiration both at the stopped-flow apparatus and at the ping-pong table. Next time I will beat you in both games.

Mr. Raymond Awad, for his friendship and for all the pizzas and Cokes that we have shared, and together with Miss Mia Jüllig, for their help in putting this thesis in printable form.

Mr. Piotr Wlad and Mrs. Åsa Genstam, for excellent technical support, and Mrs. Stina Nilsson, for exceptional secretarial work, all being extremely helpful whenever required. The staff at the library at the BMC, especially Mr. Magin Rovira, for invaluable help in retrieving references

Prof. Staffan Eriksson, for providing good working facilities, and all co-workers with which we share localities, for providing the pleasant and friendly environment that is characteristic for this department.

Finally, my family deserves my most true gratitude, for the all the care, love and support they have given me, especially my children, Sebastian and Violeta, my wife, Sissy and my mother, Lucia.

References

Abe, M., Abe, K., Kuroda, M., Arai, S. (1992) Eur. J. Biochem., 209, 933-7.

Abe, K., Emori, Y., Kondo, H., Arai, S., Suzuki, K. (1988) J. Biol. Chem., 263, 7655-9.

Abrahamson, M., Barrett, A. J., Salvesen, G., Grubb, A. (1986) J. Biol. Chem., 261, 11282-9.

Abrahamson, M., Grubb, A. (1994) Proc. Natl. Acad. Sci. U S A, 91, 1416-20.

Abrahamson, M., Mason, R. W., Hansson, H., Buttle, D. J., Grubb, A., Ohlsson, K. (1991) Biochem. J., 273, 621-6.

Abrahamson, M., Ritonja, A., Brown, M. A., Grubb, A., Machleidt, W., Barrett, A. J. (J. Biol. Chem., 262, 9688-94.

Abrahamson, M., Wikström, M., Potempa, J., Renvert, S., Hall, A. (1997) Mol. Pathol., 50, 291-7.

Agarwala, K. L., Kawabata, S., Hirata, M., Miyagi, M., Tsunasawa, S., Iwanaga, S. (1996) J. Biochem., 119, 85-94.

Alves, J. B., Alves, M. S., Naito, Y. (1994) Mem. Inst. Oswaldo Cruz, 89, 81-5.

Anastasi, A., Brown, M. A., Kembhavi, A. A., Nicklin, M. J., Sayers, C. A., Sunter, D. C., Barrett, A. J. (1983) Biochem. J., 211, 129-38.

Aoki, H., Akaike, T., Abe, K., Kuroda, M., Arai, S., Okamura, R., Negi, A., Maeda, H. (1995) Antimicrob. Agents Chemother., 39, 846-9.

Auerswald, E. A., Genenger, G., Assfalg-Machleidt, I., Machleidt, W., Engh, R. A., Fritz, H. (1992) Eur. J. Biochem., 209, 837-45.

Auerswald, E. A., Nagler, D. K., Assfalg-Machleidt, I., Stubbs, M. T., Machleidt, W., Fritz, H. (1995) FEBS Lett., 361, 179-84.

Authier, F., Posner, B. I., Bergeron, J. J. (1996) FEBS Lett., 389, 55-60.

Baker, E. N. (1980) J. Mol. Biol., 141, 441-84.

Balls, A.K., Lineweaver, H. and Thompson, R.R. (1937) Science 86, 379

Barrett, A. J. (1981) Methods. Enzymol., 771-778.

Barrett, A. J. (1986) Biomed. Biochim. Acta, 45, 1363-74.

Barrett, A. J., Davies, M. E., Grubb, A. (1984) Biochem. Biophys. Res. Commun., 120, 631-6.

Barrett, A. J., Kirschke, H. (1981) Methods. Enzymol., 535-61.

Bieth, J. G. (1995) Methods. Enzymol., 248, 59-84.

Björck, L., Grubb, A., Kjellen, L. (1990) J. Virol., 64, 941-3.

Björk, I., Ylinenjärvi, K. (1989) Biochem. J., 260, 61-8.

Björk, I., Ylinenjärvi, K. (1990) Biochemistry 29, 1770-6.

Björk, I., Alriksson, E., Ylinenjärvi, K. (1989) Biochemistry 28, 1568-73.

Björk, I., Brieditis, I., Abrahamson, M. (1995) Biochem. J., 306, 513-8.

Björk, I., Brieditis, I., Raub-Segall, E., Pol, E., Håkansson, K., Abrahamson, M. (1996) Biochemistry 35, 10720-6.

Björk, I., Pol, E., Raub-Segall, E., Abrahamson, M., Rowan, A. D., Mort, J. S. (1994) Biochem. J., 299, 219-25.

Blankenvoorde, M. F., Henskens, Y. M., van't Hof, W., Veerman, E. C., Nieuw Amerongen, A. V. (1996) Biol. Chem., 377, 847-50.

Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., Turk, V. (1988) Embo J. 7, 2593-9.

Bohley, P., Seglen, P. O. (1992) Experientia 48, 151-7.

Boller, T. (1986) In: Plant Proteolytic Enzymes, Vol. 1, (Dalling, M.J., Ed.), CRC Press, Boca Raton, pp. 67-96

Bossard, M. J., Tomaszek, T. A., Thompson, S. K., Amegadzie, B. Y., Hanning, C. R., Jones, C., Kurdyla, J. T., McNulty, D. E., Drake, F. H., Gowen, M., Levy, M. A. (1996) J. Biol. Chem., 271, 12517-24.

Brillard-Bourdet, M., Nguyen, V., Ferrer-di Martino, M., Gauthier, F., Moreau, T. (1998) Biochem. J., 331, 239-44.

Brocklehurst, K., Kowlessur, D., O'Driscoll, M., Patel, G., Quenby, S., Salih, E., Templeton, W., Thomas, E. W., Willenbrock, F. (1987) Biochem. J., 244, 173-81.

Brocklehurst, K., Willenbrock, F. and Salih, E. (1987) In: Hydrolytic enzymes, (Neurberger, A. and Brocklehurst, K., Eds.), Elsavier, Amsterdam-New York-Oxford, pp. 39-158

Brzin, J., Kopitar, M., Turk, V., Machleidt, W. (1983) Hoppe Seylers Z. Physiol. Chem., 364, 1475-80.

Buck, M. R., Karustis, D. G., Day, N. A., Honn, K. V., Sloane, B. F. (1992) Biochem. J., 282, 273-8.

Chagas, J. R., Authie, E., Serveau, C., Lalmanach, G., Juliano, L., Gauthier, F. (1997) Mol. Biochem. Parasitol., 88, 85-94.

Chauhan, S. S., Goldstein, L. J., Gottesman, M. M. (1991) Cancer Res., 51, 1478-81.

Collins, A. R., Grubb, A. (1998) Oral Microbiol. Immunol., 13, 59-61.

Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., Mort, J. S., Cygler, M. (1996) Embo J., 15, 5492-503.

Del Nery, E., Juliano, M. A., Lima, A. P., Scharfstein, J., Juliano, L. (1997) J. Biol. Chem., 272, 25713-8.

DeLa Cadena, R. A., Colman, R. W. (1991) Trends Pharmacol. Sci., 12, 272-5.

Dieckmann, T., Mitschang, L., Hofmann, M., Kos, J., Turk, V., Auerswald, E. A., Jaenicke, R., Oschkinat, H. (1993) J. Mol. Biol., 234, 1048-59.

Drake, F. H., Dodds, R. A., James, I. E., Connor, J. R., Debouck, C., Richardson, S., Lee-Rykaczewski, E., Coleman, L., Rieman, D., Barthlow, R., Hastings, G., Gowen, M. (1996) J. Biol. Chem., 271, 12511-6.

Drenth, J., Kalk, K. H., Swen, H. M. (1976) Biochemistry 15, 3731-8.

Dunn, A. D., Myers, H. E., Dunn, J. T. (1996) Endocrinology 137, 3279-85.

Ebert, W., Knoch, H., Werle, B., Trefz, G., Muley, T., Spiess, E. (1994) Anticancer Res., 14, 895-9.

Ebert, E., Werle, B., Julke, B., Kopitar-Jerala, N., Kos, J., Lah, T., Abrahamson, M., Spiess, E., Ebert, W. (1997) Adv. Exp. Med. Biol., 421, 259-65.

Ekiel, I., Abrahamson, M., Fulton, D. B., Lindahl, P., Storer, A. C., Levadoux, W., Lafrance, M., Labelle, S., Pomerleau, Y., Groleau, D., LeSauteur, L., Gehring, K. (1997) J. Mol. Biol., 271, 266-77.

Enenkel, C., Wolf, D. H. (1993) J. Biol. Chem., 268, 7036-43.

Esnard, F., Gauthier, F. (1983) J. Biol. Chem., 258, 12443-7.

Esnard, A., Esnard, F., Faucher, D., Gauthier, F. (1988) FEBS Lett., 236, 475-8.

Estrada, S., Nycander, M., Hill, N. J., Craven, C. J., Waltho, J. P., Björk, I. (1998) Biochemistry 37, 7551-60.

Evans, H. J., Barrett, A. J. (1987) Biochem. J., 246, 795-7.

Fong, D., Kartasova, T., Sloane, B. F., Chan, M. M. (1989) FEBS Lett., 257, 55-8.

Fossum, K., Whitaker, J. R. (1968) Arch. Biochem. Biophys., 125, 367-75.

Freije, J. P., Abrahamson, M., Olafsson, I., Velasco, G., Grubb, A., Lopez-Otin, C. (1991) J. Biol. Chem., 266, 20538-43.

Gelb, B. D., Shi, G. P., Chapman, H. A., Desnick, R. J. (1996) Science 273, 1236-8.

Genenger, G., Lenzen, S., Mentele, R., Assfalg-Machleidt, I., Auerswald, E. A. (1991) Biomed. Biochim. Acta, 50, 621-5.

Green, G. D., Kembhavi, A. A., Davies, M. E., Barrett, A. J. (1984) Biochem. J., 218, 939-46.

Grenier, D. (1996) Microbiology 142, 955-61.

Grubb, A., Löfberg, H. (1982) Proc. Natl. Acad. Sci. U S A, 79, 3024-7.

Guncar, G., Podobnik, M., Pungercar, J., Strukelj, B., Turk, V., Turk, D. (1998) Structure 6, 51-61.

Hall, A., Dalbøge, H., Grubb, A., Abrahamson, M. (1993) Biochem. J., 291, 123-9.

Hall, A., Ekiel, I., Mason, R. W., Kasprzykowski, F., Grubb, A., Abrahamson, M. (1998) Biochemistry 37, 4071-9.

Hall, A., Håkansson, K., Mason, R. W., Grubb, A., Abrahamson, M. (1995) J. Biol. Chem., 270, 5115-21.

Henskens, Y. M., Van der Velden, U., Veerman, E. C., Nieuw Amerongen, A. V. (1993) Ann. N. Y. Acad. Sci., 694, 280-2.

Henskens, Y. M., Veerman, E. C., Mantel, M. S., van der Velden, U., Nieuw Amerongen, A. V. (1994) J. Dent. Res., 73, 1606-14.

Henskens, Y. M., Veerman, E. C., Nieuw Amerongen, A. V. (1996) Biol. Chem. Hoppe Seyler., 377, 71-86.

Hirado, M., Iwata, D., Niinobe, M., Fujii, S. (1981) Biochim. Biophys. Acta, 669, 21-7.

Hirado, M., Tsunasawa, S., Sakiyama, F., Niinobe, M., Fujii, S. (1985) FEBS Lett., 186, 41-5.

Hopsu-Havu, V. K., Järvinen, M., Rinne, A. (1983) Br. J. Dermatol., 25, 77-85.

Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T., Mort, J. S. (1997) J. Biol. Chem., 272, 1197-202.

Inoue, T., Ishida, T., Sugio, K., Sugimachi, K. (1994) Cancer Res., 54, 6133-6.

Inui, T., Ishibashi, O., Inaoka, T., Origane, Y., Kumegawa, M., Kokubo, T., Yamamura, T. (1997) J. Biol. Chem., 272, 8109-12.

Irvine, J. W., Coombs, G. H., North, M. J. (1992) Fems. Microbiol. Lett., 75, 67-72.

Isemura, S., Saitoh, E., Ito, S., Isemura, M., Sanada, K. (1984) J. Biochem., 96, 1311-4.

Isemura, S., Saitoh, E., Sanada, K. (1986) FEBS Lett., 198, 145-9.

Isemura, S., Saitoh, E., Sanada, K. (1987) J. Biochem., 102, 693-704.

Iwata, Y., Mort, J. S., Tateishi, H., Lee, E. R. (1997) Arthritis Rheum.,., 40, 499-509.

Järvinen, M. (1976) Acta. Chem. Scand.,[b] 30, 933-40.

Järvinen, M. (1978) J. Invest. Dermatol., 71, 114-8.

Järvinen, M., Rinne, A. (1982) Biochim. Biophys. Acta, 708, 210-7.

Järvinen, M., Rinne, A., Hopsu-Havu, V. K. (1987) Acta. Histochem., 82, 5-18.

Jensson, O., Gudmundsson, G., Arnason, A., Blöndal, H., Petursdottir, I., Thorsteinsson, L., Grubb, A., Löfberg, H., Cohen, D., Frangione, B. (1987) Acta. Neurol. Scand., 76, 102-14.

Jerala, R., Kroon-Zitko, L., Turk, V. (1994) Protein Expr. Purif., 5, 65-9.

Jerala, R., Trstenjak-Prebanda, M., Kroon-Zitko, L., Lenarcic, B., Turk, V. (1990) Biol. Chem. Hoppe Seyler., 371, 157-60.

Kafienah, W., Brömme, D., Buttle, D. J., Croucher, L. J., Hollander, A. P. (1998) Biochem. J., 331, 727-32.

Kaji, H., Kumagai, I., Takeda, A., Miura, K., Samejima, T. (1989) J. Biochem., 105, 143-7.

Kamphuis, I. G., Drenth, J., Baker, E. N. (1985) J. Mol. Biol., 182, 317-29.

Kamphuis, I. G., Kalk, K. H., Swarte, M. B., Drenth, J. (1984) J. Mol. Biol., 179, 233-56.

Keilova, H., Tomasek, V. (1977) Acta Biol. Med. Ger., 36, 1873-81.

Kimura, M., Ikeda, T., Fukumoto, D., Yamasaki, N., Yonekura, M. (1995) Biosci. Biotechnol. Biochem., 59, 2328-9.

Kirschke, H., Wiederanders, B. (1994) Methods. Enzymol., 244, 500-11.

Kirschke, H., Barrett, A. J., Rawlings, N. D. (1995) Protein Profile 2, 1581-643.

Kirschke, H., Wiederanders, B., Brömme, D., Rinne, A. (1989) Biochem. J., 264, 467-73.

Koide, Y., Noso, T. (1994) Biosci. Biotechnol. Biochem., 58, 164-9.

Kondo, H., Abe, K., Nishimura, I., Watanabe, H., Emori, Y., Arai, S. (1990) J. Biol. Chem., 265, 15832-7.

Krepela, E., Kasafirek, E., Novak, K., Viklicky, J. (1990) Neoplasma 37, 61-70.

Krizaj, I., Turk, B., Turk, V. (1992) FEBS Lett., 298, 237-9.

Lafreniere, R. G., Rochefort, D. L., Chretien, N., Rommens, J. M., Cochius, J. I., Kalviainen, R., Nousiainen, U., Patry, G., Farrell, K., Söderfeldt, B., Federico, A., Hale, B. R., Cossio, O. H., Sörensen, T., Pouliot, M. A., Kmiec, T., Uldall, P., Janszky, J., Pranzatelli, M. R., Andermann, F., Andermann, E., Rouleau, G. A. (1997) Nat. Genet., 15, 298-302.

Lah, T. T., Buck, M. R., Honn, K. V., Crissman, J. D., Rao, N. C., Liotta, L. A., Sloane, B. F. (1989) Clin. Exp. Metastasis, 7, 461-8.

Lah, T. T., Clifford, J. L., Helmer, K. M., Day, N. A., Moin, K., Honn, K. V., Crissman, J. D., Sloane, B. F. (1989) Biochim. Biophys. Acta, 993, 63-73.

Lah, T. T., Kokalj-Kunovar, M., Drobnic-Kosorok, M., Babnik, J., Golouh, R., Vrhovec, I., Turk, V. (1992) Biol. Chem. Hoppe Seyler., 373, 595-604.

Lalioti, M. D., Mirotsou, M., Buresi, C., Peitsch, M. C., Rossier, C., Ouazzani, R., Baldy-Moulinier, M., Bottani, A., Malafosse, A., Antonarakis, S. E. (1997) Am. J. Hum. Genet., 60, 342-51.

Lalioti, M. D., Scott, H. S., Buresi, C., Rossier, C., Bottani, A., Morris, M. A., Malafosse, A., Antonarakis, S. E. (1997) Nature 386, 847-51.

Lemaire, R., Flipo, R. M., Migaud, H., Fontaine, C., Huet, G., Dacquembronne, E., Lafyatis, R. (1997) Arthritis Rheum., 40, 1540-2.

Lenarcic, B., Gabrijelcic, D., Rozman, B., Drobnic-Kosorok, M., Turk, V. (1988) Biol. Chem. Hoppe Seyler., 369, 257-61.

Lenarcic, B., Krizaj, I., Zunec, P., Turk, V. (1996) FEBS Lett., 395, 113-8.

Lerner, U. H., Johansson, L., Ranjsö, M., Rosenquist, J. B., Reinholt, F. P., Grubb, A. (1997) Acta Physiol. Scand., 161, 81-92.

Lindahl, P., Abrahamson, M., Björk, I. (1992) Biochem. J., 281, 49-55.

Lindahl, P., Alriksson, E., Jornvall, H., Björk, I. (1988) Biochemistry 27, 5074-82.

Lindahl, P., Nycander, M., Ylinenjärvi, K., Pol, E., Björk, I. (1992) Biochem. J., 286, 165-71.

Lindahl, P., Ripoll, D., Abrahamson, M., Mort, J. S., Storer, A. C. (1994) Biochemistry 33, 4384-92.

Machleidt, W., Borchart, U., Fritz, H., Brzin, J., Ritonja, A., Turk, V. (1983) Hoppe Seylers Z. Physiol. Chem., 364, 1481-6.

Machleidt, W., Thiele, U., Assfalg-Machleidt, I., Forger, D., Auerswald, E. A. (1991) Biomed. Biochim. Acta, 50, 613-20.

Machleidt, W., Thiele, U., Laber, B., Assfalg-Machleidt, I., Esterl, A., Wiegand, G., Kos, J., Turk, V., Bode, W. (1989) FEBS Lett., 243, 234-8.

Machleidt, W., Nagler, D. K., Assfalg-Machleidt, I., Stubbs, M. T., Fritz, H., Auerswald, E. A. (1995) FEBS Lett., 361, 185-90.

Mackay, E. A., Ehrhard, A., Moniatte, M., Guenet, C., Tardif, C., Tarnus, C., Sorokine, O., Heintzelmann, B., Nay, C., Remy, J. M., Higaki, J., Van Dorsselaer, A., Wagner, J., Danzin, C., Mamont, P. (1997) Eur. J. Biochem., 244, 414-25.

Maes, D., Bouckaert, J., Poortmans, F., Wyns, L., Looze, Y. (1996) Biochemistry 35, 16292-8.

Mantle, D., Falkous, G., Ishiura, S., Perry, R. H., Perry, E. K. (1995) J. Neurol. Sci., 131, 65-70.

Martin, J. R., Craven, C. J., Jerala, R., Kroon-Zitko, L., Zerovnik, E., Turk, V., Waltho, J. P. (1995) J. Mol. Biol., 246, 331-43.

Mason, R. W., Green, G. D., Barrett, A. J. (1985) Biochem. J., 226, 233-41.

Mason, R. W., Sol-Church, K., Abrahamson, M. (1998) Biochem. J., 330, 833-8.

McGrath, M. E., Eakin, A. E., Engel, J. C., McKerrow, J. H., Craik, C. S., Fletterick, R. J. (1995) J. Mol. Biol., 247, 251-9.

Misaka, T., Kuroda, M., Iwabuchi, K., Abe, K., Arai, S. (1996) Eur. J. Biochem., 240, 609-14.

Mizuochi, T., Yee, S. T., Kasai, M., Kakiuchi, T., Muno, D., Kominami, E. (1994) Immunol. Lett., 43, 189-93.

Müller-Esterl, W., Lottspeich, F., Kellermann, J., Henschen, A. (1986) Behring. Inst. Mitt., 48-61.

Murzin, A. G. (1993) J. Mol. Biol., 230, 689-94.

Musil, D., Zucic, D., Turk, D., Engh, R. A., Mayr, I., Huber, R., Popovic, T., Turk, V., Towatari, T., Katunuma, N., et, a. (1991) Embo J., 10, 2321-30.

Naito, Y., Suzuki, I., Hasegawa, S. (1992) Comp. Biochem. Physiol., [b] 102, 861-5.

Naito, Y., Sasaki, M., Umemoto, T., Namikawa, I., Sakae, K., Ishihara, Y., Isomura, S., Suzuki, I. (1995) Comp. Biochem. Physiol. C. Pharmacol. Toxicol. Endocrinol., 110, 71-5.

Neves, F. A., Duncan, K. G., Baxter, J. D. (1996) Hypertension 27, 514-7.

Ni, J., Abrahamson, M., Zhang, M., Fernandez, M. A., Grubb, A., Su, J., Yu, G. L., Li, Y., Parmelee, D., Xing, L., Coleman, T. A., Gentz, S., Thotakura, R., Nguyen, N., Hesselberg, M., Gentz, R. (1997) J. Biol. Chem., 272, 10853-8.

Nicklin, M. J., Barrett, A. J. (1984) Biochem. J., 223, 245-53.

Nikawa, T., Towatari, T., Ike, Y., Katunuma, N. (1989) FEBS Lett., 255, 309-14.

North, M. J. (1992) Biol. Chem. Hoppe Seyler., 373, 401-6.

Nycander, M., Björk, I. (1990) Biochem. J., 271, 281-4.

Nycander, M., Estrada, S., Mort, J. S., Abrahamson, M., Björk, I. (1998) FEBS Lett., 422, 61-4.

O'Hara, B. P., Hemmings, A. M., Buttle, D. J., Pearl, L. H. (1995) Biochemistry 34, 13190-5.

Okada, Y., Teno, N., Tsuboi, S., Nakabayashi, K., Itoh, N., Okamoto, H., Nishi, N. (1988) Chem. Pharm. Bull., 36, 1982-9.

Okada, Y., Tsuda, Y., Teno, N., Wanaka, K., Bohgaki, M., Hijikata-Okunomiya, A., Naito, T., Okamoto, S. (1988) Chem. Pharm. Bull., 36, 1289-97.

Olsson, S. L., Ek, B., Wilm, M., Broberg, S., Rask, L., Björk, I. (1997) Biochim. Biophys. Acta, 1343, 203-10.

Palsdottir, A., Abrahamson, M., Thorsteinsson, L., Arnason, A., Olafsson, I., Grubb, A., Jensson, O. (1988) Lancet 2, 603-4.

Pennacchio, L. A., Myers, R. M. (1996) Genome Res 6, 1103-9.

Pennacchio, L. A., Lehesjoki, A. E., Stone, N. E., Willour, V. L., Virtaneva, K., Miao, J., D'Amato, E., Ramirez, L., Faham, M., Koskiniemi, M., Warrington, J. A., Norio, R., de la Chapelle, A., Cox, D. R., Myers, R. M. (1996) Science 271, 1731-4.

Pickersgill, R.W., Razkallah, P., Harris, W.G. and Goodenough, P.W. (1991) Acta Cryst. B 47, 766-771

Pol, E., Olsson, S. L., Estrada, S., Prasthofer, T. W., Björk, I. (1995) Biochem. J., 311, 275-82.

Poulik, M. D., Shinnick, C. S., Smithies, O. (1981) Mol Immunol 18, 569-72.

Qian, F., Bajkowski, A. S., Steiner, D. F., Chan, S. J., Frankfater, A. (1989) Cancer Res., 49, 4870-5.

Rawlings, N. D., Barrett, A. J. (1994) Methods. Enzymol., 244, 461-86.

Rawlings, N. D., Barrett, A. J. (1998) MEROPS, http://www.bi.bbsrc.ac.uk/Merops/Merops.htm

Rawlings, N. D., Pearl, L. H., Buttle, D. J. (1992) Biol. Chem. Hoppe Seyler., 373, 1211-5.

Rempel, S. A., Rosenblum, M. L., Mikkelsen, T., Yan, P. S., Ellis, K. D., Golembieski, W. A., Sameni, M., Rozhin, J., Ziegler, G., Sloane, B. F. (1994) Cancer Res., 54, 6027-31.

Rozhin, J., Wade, R. L., Honn, K. V., Sloane, B. F. (1989) Biochem. Biophys. Res. Commun., 164, 556-61.

Salvesen, G., Parkes, C., Abrahamson, M., Grubb, A., Barrett, A. J. (1986) Biochem. J., 234, 429-34.

Schechter, I., Berger, A. (1967) Biochem. Biophys. Res. Commun., 27, 157-62.

Scholze, H., Tannich, E. (1994) Methods. Enzymol., 244, 512-23.

Schwabe, C., Anastasi, A., Crow, H., McDonald, J. K., Barrett, A. J. (1984) Biochem. J., 217, 813-7.

Sen, L. C., Whitaker, J. R. (1973) Arch. Biochem. Biophys., 158, 623-32.

Shibuya, K., Kaji, H., Itoh, T., Ohyama, Y., Tsujikami, A., Tate, S., Takeda, A., Kumagai, I., Hirao, I., Miura, K., et, a. (1995) Biochemistry 34, 12185-92.

Shibuya, K., Kaji, H., Ohyama, Y., Tate, S., Kainosho, M., Inagaki, F., Samejima, T. (1995) J. Biochem., 118, 635-42.

Shiraishi, T., Mori, M., Tanaka, S., Sugimachi, K., Akiyoshi, T. (1998) Int. J. Cancer, 79, 175-8.

Sivaparvathi, M., Sawaya, R., Wang, S. W., Rayford, A., Yamamoto, M., Liotta, L. A., Nicolson, G. L., Rao, J. S. (1995) Clin. Exp. Metastasis, 13, 49-56.

Sloane, B. F., Rozhin, J., Hatfield, J. S., Crissman, J. D., Honn, K. V. (1987) Exp. Cell. Biol., 55, 209-24.

Solem, M., Rawson, C., Lindburg, K., Barnes, D. (1990) Biochem. Biophys. Res. Commun., 172, 945-51.

Song, I., Taylor, M., Baker, K., Bateman, R., Jr. (1995) Gene 162, 221-4.

Sotiropoulou, G., Anisowicz, A., Sager, R. (1997) J. Biol. Chem., 272, 903-10.

Stoka, V., Nycander, M., Lenarcic, B., Labriola, C., Cazzulo, J. J., Björk, I., Turk, V. (1995) FEBS Lett., 370, 101-4.

Storer, A. C., Menard, R. (1994) Methods. Enzymol., 244, 486-500.

Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B., Turk, V. (1990) Embo J., 9, 1939-47.

Takeda, A., Kaji, H., Nakaya, K., Aoki, Y., Nakamura, Y., Samejima, T. (1985) Biochem. Int., 11, 557-64.

Takeda, A., Kobayashi, S., Samejima, T. (1983) J. Biochem., 94, 811-20.

Takio, K., Kominami, E., Bando, Y., Katunuma, N., Titani, K. (1984) Biochem. Biophys. Res. Commun., 121, 149-54.

Takio, K., Kominami, E., Wakamatsu, N., Katunuma, N., Titani, K. (1983) Biochem. Biophys. Res. Commun., 115, 902-8.

Tate, S., Ushioda, T., Utsunomiya-Tate, N., Shibuya, K., Ohyama, Y., Nakano, Y., Kaji, H., Inagaki, F., Samejima, T., Kainosho, M. (1995) Biochemistry 34, 14637-48.

Thiele, U., Assfalg-Machleidt, I., Machleidt, W., Auerswald, E. A. (1990) Biol. Chem. Hoppe Seyler., 371, 125-36.

Torti, S. V., Torti, F. M. (1998) J. Biol. Chem., 273, 13630-5.

Troeberg, L., Pike, R. N., Morty, R. E., Berry, R. K., Coetzer, T. H., Lonsdale-Eccles, J. D. (1996) Eur. J. Biochem., 238, 728-36.

Tsai, Y. J., Chang, G. D., Huang, C. J., Chang, Y. S., Huang, F. L. (1996) Comp. Biochem. Physiol. B. Biochem. Mol. Biol., 113, 573-80.

Tsui, F. W., Tsui, H. W., Mok, S., Mlinaric, I., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Siminovitch, K. A. (1993) Genomics 15, 507-14.

Tsushima, H., Mine, H., Hoshika, K., Kawakami, Y., Hyodoh, F., Ueki, A. (1992) J. Bacteriol., 174, 4807-10.

Tsushima, H., Mine, H., Kawakami, Y., Hyodoh, F., Ueki, A. (1994) Microbiology 140, 167-71.

Tu, G. F., Cole, T., Southwell, B. R., Schreiber, G. (1990) Brain Res. Dev. Brain Res., 55, 203-8.

Turk, B., Bieth, J. G., Björk, I., Dolenc, I., Turk, D., Cimerman, N., Kos, J., Colic, A., Stoka, V., Turk, V. (1995) Biol. Chem. Hoppe Seyler., 376, 225-30.

Turk, B., Colic, A., Stoka, V., Turk, V. (1994) FEBS Lett., 339, 155-9.

Turk, B., Ritonja, A., Björk, I., Stoka, V., Dolenc, I., Turk, V. (1995) FEBS Lett., 360, 101-5.

Turk, B., Stoka, V., Björk, I., Boudier, C., Johansson, G., Dolenc, I., Colic, A., Bieth, J. G., Turk, V. (1995) Protein Sci., 4, 1874-80.

Turk, B., Turk, V., Turk, D. (1997) Biol. Chem., 378, 141-50.

Turk, D., Guncar, G., Podobnik, M., Turk, B. (1998) Biol. Chem., 379, 137-47.

Turk, V., Brzin, J., Kotnik, M., Lenarcic, B., Popovic, T., Ritonja, A., Trstenjak, M., Begic-Odobasic, L., Machleidt, W. (1986) Biomed. Biochim. Acta, 45, 1375-84.

Turk, V., Brzin, J., Longer, M., Ritonja, A., Eropkin, M., Borchart, U., Machleidt, W. (1983) Hoppe Seylers Z. Physiol. Chem., 364, 1487-96.

Umemoto, T., Naito, Y., Li, M., Suzuki, I., Namikawa, I. (1996) Lett. Appl. Microbiol., 23, 151-3.

Urwin, P. E., Atkinson, H. J., McPherson, M. J. (1995) Protein Eng., 8, 1303-7.

Virtaneva, K., D'Amato, E., Miao, J., Koskiniemi, M., Norio, R., Avanzini, G., Franceschetti, S., Michelucci, R., Tassinari, C. A., Omer, S., Pennacchio, L. A., Myers, R. M., Dieguez-Lucena, J. L., Krahe, R., de la Chapelle, A., Lehesjoki, A. E. (1997) Nat. Genet., 15, 393-6.

Werle, B., Ebert, W., Klein, W., Spiess, E. (1994) Anticancer Res., 14, 1169-76.

Willstätter, R and Bamann, E. (1929) Hoppe-Seylers Z. Physiol. Chem. 180, 127

Xing, R., Addington, A. K., Mason, R. W. (1998) Biochem. J., 332, 499-505.

Yan, S., Sameni, M., Sloane, B. F. (1998) Biol. Chem., 379, 113-23.

Acta Universitatis Agriculturae Sueciae
presents doctoral theses from the faculty of agriculture, landscape architecture and horticulture, the faculty of forestry and the faculty of veterinary medicine at the Swedish University of Agricultural Sciences. Each faculty has its own subseries, Agraria, Silvestria and Veterinaria, respectively, with separate numbering.

ISSN 1401-6257
ISBN 91-576-5448-4