Acta Biochimica Polonica

Vol. 48 No. 1/2001

1–20

QUARTERLY

Review

Structural studies of cysteine proteases and their inhibitors★✪

Zbigniew Grzonka[1]✉, Elżbieta Jankowska[1], Franciszek Kasprzykowski[1], Regina Kasprzykowska[1], Leszek Łankiewicz[1], Wiesław Wiczk[1], Ewa Wieczerzak[1], Jerzy Ciarkowski[1], Piotr Drabik[1], Robert Janowski[2], Maciej Kozak[3], Mariusz Jaskólski[2,4] and Anders Grubb[5]

[1]*Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland;* [2]*Faculty of Chemistry, A. Mickiewicz University, Poznań, Poland;* [3]*Faculty of Physics, A. Mickiewicz University, Poznań, Poland;* [4]*Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland;* [5]*Department of Clinical Chemistry, University Hospital, Lund, Sweden*

Received: 12 December, 2000; accepted: 29 January, 2001

Key words: cysteine proteases, cystatins, synthetic inhibitors, structure–activity relationship

Cysteine proteases (CPs) are responsible for many biochemical processes occurring in living organisms and they have been implicated in the development and progression of several diseases that involve abnormal protein turnover. The activity of CPs is regulated among others by their specific inhibitors: cystatins. The main aim of this review is to discuss the structure–activity relationships of cysteine proteases and cystatins, as well as of some synthetic inhibitors of cysteine proteases structurally based on the binding fragments of cystatins.

CYSTEINE PROTEASES (CPs)

Proteases are classified according to their catalytic site into four major classes: serine proteases, cysteine proteases, aspartic proteases and metallo-proteases [1, 2]. Cysteine proteases (CPs) are proteins with molecular mass about 21–30 kDa. They show the highest hydrolytic activity at pH 4–6.5. Because of the high tendency of the thiol group to oxidation, the environment of the

★Presented at the International Conference on "Conformation of Peptides, Proteins and Nucleic Acids", Debrzyno, Poland, 2000.

✪This work was supported by the State Committee for Scientific Research (KBN, Poland) grant No. 279/P04/97/13.

✉Corresponding author: Prof. Zbigniew Grzonka, Department of Organic Chemistry, Faculty of Chemistry, University of Gdańsk, J. Sobieskiego 18, 80-952 Gdańsk, Poland; tel.: (48 58) 345 0369; fax.: (48 58) 344 9680; e-mail: grzonka@chemik.chem.univ.gda.pl

Abbreviations: Boc, *t*-butoxycarbonyl; CP, cysteine protease; DAM, diazomethylketone; dArg, desaminearginine; E64, (2*S*,3*S*)-*trans*-epoxysuccinyl-L-leucyl-agmatine; hCA, hCB, hCC, hCD, hCE, hCF, hCS, hCSA, hCSN, human cystatin A, B, C, D, E, F, S, SA, SN, respectively; hHK, hLK, human high or low molecular mass kininogen, respectively; ValΨ[CH_2NH], (*S*)-1-isopropyl-1,2-etanediamine; Z, benzyloxycarbonyl.

enzyme should contain a reducing component. Glutathione serves as an activating agent in cells, whereas addition of mercaptoethanol or dithiothreitol is required for *in vitro* experiments.

Cysteine proteases are present in all living organisms. Till now, 21 families of CPs have been discovered [3, 4], almost half of them in viruses. Many of these enzymes are found in bacteria (e.g. clostripain in *Clostridium histolyticum,* gingipain in *Porhyromonas gingivalis*), fungi (cathepsin B in *Aspergillus flavus,* proteinease ysc F in yeast), protozoa (cruzipain in *Trypanosoma cruzi,* amoebopain in *Entamoeba histolytica*) and plants (papain and chymopapain in *Carica papaya,* ficin in *Ficus glabrata,* bromelain in *Ananas comosus,* actinidin in *Actinidia chinesis,* calotropin DI in *Calotropis gigantea*). In mammals, two main groups of cysteine proteases are present: cytosolic calpains (calpain type I, calpain type II) and lysosomal cathepsins (cathepsins: B, C, H, K, L, M, N, S, T, V, and W) [3–5].

The best characterized family of cysteine proteases is that of papain. The papain family contains peptidases which are structurally related to papain, like for example lysosomal cathepsins. Papain is characterized by a two-domain structure (Fig. 1). The active site (catalytic pocket), where

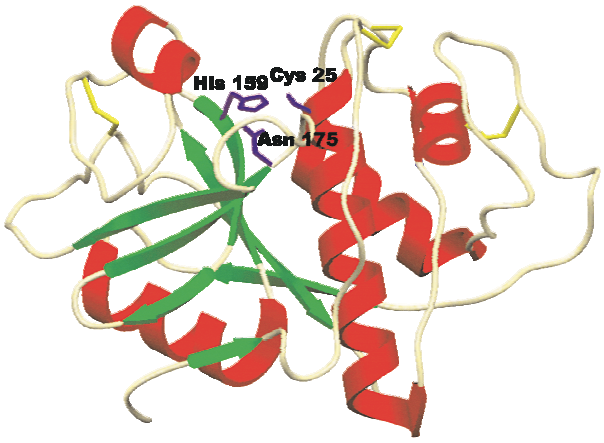

Figure 1. Structure of papain [100].

the substrate is bound, is located between the domains. The catalytic residues of papain are Cys^{25} and His^{159}, and they are evolutionarily preserved in all CPs. Following a proposal by Schechter & Berger [6], the substrate pocket of papain binds at least seven amino-acid residues in appropriate subsites (Fig. 2). Recently, Turk *et al.* [7] have proposed, on the basis of kinetic and structural studies, that only 5 subsites are important for substrate binding. The S_2, S_1, and S_1' pockets are important for both backbone and side-chain binding, whereas S_3 and S_2' are crucial only for amino acid side-chain binding.

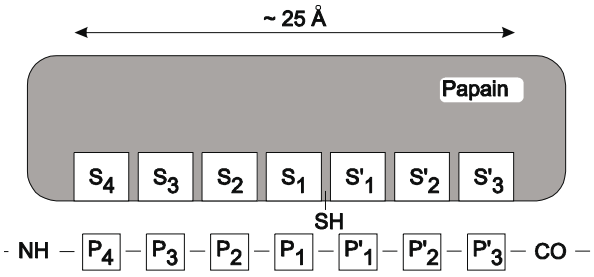

Figure 2. Substrate subsites of papain [6].

Enzymatic activity of cysteine proteases is related to the presence of a catalytic diad formed by the cysteine and histidine residues which in the pH interval 3.5–8.0, exists as an ion-pair **–S^-...H^+Im–** [8, 9]. Formation of an intermediate, S-acyl-enzyme moiety, is a fundamental step in hydrolysis. This intermediate is formed *via* nucleophilic attack of the thiolate group of the cysteine residue on the carbonyl group of the hydrolyzed peptide bond with a release of the C-terminal fragment of the cleaved product. In the next step, a water molecule reacts with the intermediate, the N-terminal fragment is released, and the regenerated free papain molecule can begin a new catalytic cycle [10].

CPs are responsible for many biochemical processes occurring in living organisms. The main physiological role of CPs is metabolic degradation of peptides and proteins. Mammalian cysteine proteases have been implicated in the development and progression of many diseases that involve abnormal protein turnover [11–15]. The activity of cysteine proteases is regulated by proper gene transcription and the rate of protease synthesis and degradation, as well as by their specific inhibitors.

CYSTATINS

Many natural protein inhibitors of cysteine proteases, called cystatins, have been isolated and characterized. They act both intra- and extracellularly forming complexes with their target enzymes. Maintenance of appropriate equilibrium between free cysteine proteases and their complexes with inhibitors is critical for proper functioning of all living systems. In this role, cystatins are general regulators of harmful cysteine protease activities. The roles of cystatins in health and disease have been reviewed by Henskens *et al.* [13] and Grubb [14].

The human superfamily of cystatins is divided into three families. Family I, called stefins, comprises intracellular cystatins A and B. Family II includes extracellular and/or transcellular cystatins (cystatins: C, D, E, F, S, SA, and SN). Kininogens, the intravascular cystatins, form family III of cystatins.

Stefin family

The stefin family comprises the following inhibitors: human stefin A (hCA), human stefin B (hCB) [16–18], as well as their analogues from rat [19], bovine [20, 21] and porcine [22] tissues, and from some plants [23]. Stefins are proteins of about 100 amino-acid residues (molecular mass about 11 kDa) which do not contain any sugar moiety or disulfide bridge. There is one cysteine residue in the stefin B sequence, and it can be converted into an intermolecular disulfide bridge ($-Cys^{3}-Cys^{3'}-$) resulting in formation of inactive dimers, easily transformed back into active monomers at reducing conditions [24, 25]. Immunohistochemical studies have shown the presence of stefin A in skin and epithelium, suggesting that the major function of stefin A is related to protection of these organs against overreactivity of cysteine proteases [26]. On the other hand, stefin B is distributed in many tissues, which suggests that this inhibitor interacts with cathepsins liberated from lysosomes [27]. Both inhibitors have been found in all human fluids, but at a small concentration [28]. The gene coding for human cystatin A has been assigned to chromosome 3 [29], and that for human stefin B to chromosome 21 [30].

Cystatin family

The cystatin family comprises the following human cystatins: C (hCC), D (hCD), E (hCE), F (hCF), S (hCS), SA (hCSA) and SN (hCSN). Their homologues have also been found in other mammalian organisms and birds. Chicken cystatin has been used in defining the superfamily of cystatins [31, 32]. Human cystatins are coded on chromosome 20 [33]. They consist of 120–122 amino-acid residues and are synthesized as proproteins containing a signal peptide (20 residues), which suggests that cystatins display an extracellular activity [34]. The cystatins contain two disulfide bridges and most of them are not glycosylated. Cystatins S, SA, and SN (S-type cystatins) consisting of 121 amino-acid residues with molecular mass of 14.2–14.4 kDa display high sequence homology (90%). Post-translational phosphorylation of cystatins S and SA leads to formation of several isoforms. Expression of these cystatins is very restricted: cystatin SN has been found only in saliva and tears, whereas variants S and SA are also present in seminal fluid [35]. Cystatin D has also been found mostly in saliva and tears. Fully active cystatin D, formed after removal of the 20-residue signal peptide, consists of 122 amino-acid residues with molecular mass about 13.8 kDa. The protein exists in 2 polymorphic forms: [Cys^{26}]hCD and [Arg^{26}]hCD which have identical activity, stability, and distribution. The quite low homology with other cystatins (51–55%) suggests that, on a phylogenetic tree, cystatin D is located between cystatins S and C [36, 37]. Human cystatin E (hCE, [38]), also described in the literature as cystatin M [39], and human cystatin F (hCF, [40]), also called leukocystatin [41], are released from appropriate proproteins containing signal peptides (28-mer for hCE and 19-mer for hCF). Cystatins E and F are glycoproteins built of 122 and 126 amino-acid residues, respectively. Their structures display low homology to the second family of cystatins: 26–34% in a case of hCE and 30–34% in a case of hCF. Unlike other mem-

bers of the second family, cystatin F contains an additional, third, disulfide bridge stabilizing the N-terminal fragment of the protein. Tissue-distribution profile studies have shown that the highest concentration of hCE is in the uterus and liver [38] and that of hCF in spleen and leukocytes [40].

Human cystatin C (hCC), formed after removal of a 26-residue signal peptide, is a protein of 120 amino-acid residues with molecular mass of 13.4 kDa [42, 43]. In contrast to other members of the family, hCC is a basic protein (pI = 9.3) [44]. It is widely distributed in all physiological fluids. The highest amounts were found in seminal plasma and cerebro-spinal fluid, and much lower concentrations were observed in tears, amniotic fluid, saliva, milk, and blood plasma [45]. The wide distribution and high inhibitory potency of cystatin C suggest that this protein is a major cysteine protease inhibitor.

Kininogens

The third family consists of 3 members: human high molecular mass kininogen (hHK), about 120 kDa; human low molecular mass kininogen (hLK), about 68 kDa; and kininogen T, discovered so far only in rats [46]. Kininogens hHK and hLK are glycoproteins released as proproteins containing a signal peptide (18 amino-acid residues) [3]. The highest concentration of kininogens is found in blood plasma and synovial fluid [45].

Other cystatins and cystatin-like proteins

A number of proteins have been described which, in spite of high sequence homology, show distinct differences in structure and biological activity in comparison with cystatins. Histidine-rich glycoproteins (HRG) and fetuins are examples of such cystatin-like proteins. Both HRG and fetuins did not inhibit cysteine proteases. The subject of cystatins and cystatin-like proteins has been reviewed by Brown & Dziegielewska [47]. Conversely, there are also proteins, like the intensely sweet plant protein monellin which, in spite of very low sequence homology and lack of inhibitory function, have a cystatin-like three-dimensional structure [48, 49].

Cysteine protease–cystatin interaction

Numerous spectroscopic, kinetic, and crystallographic studies have been carried out to explain the mechanism of cysteine protease inhibition by cystatins. The results have shown that the inhibitor binds in a one-step process that is simple, reversible, and second-order type. In addition, those studies have revealed that enzymes with a blocked active centre could still bind cystatins, albeit with lower affinity [50–52]. This indicates that cysteine protease–cystatin interactions are not based on a simple reaction with the catalytic cysteine residue of the enzyme, as is typical of substrates, but that they consist of hydrophobic contacts between the binding regions of cystatins and the corresponding residues forming the binding pockets of the enzyme. Despite their structural homology and similar mode of inhibition, cystatins display quite different enzyme affinities (Table 1).

From functional studies of cystatin C it was concluded that the N-terminal fragment containing 11 amino-acid residues is important for the inhibitory activity of hCC [44]. Our early studies with synthetic peptides corresponding to the N-terminal sequence of hCC showed that they were very good substrates of papain, and that the cleavage took place at the Gly^{11}–Gly^{12} peptide bond. We have concluded that Arg^{8}, Leu^{9} and Val^{10} from the N-terminal segment of cystatin C interact with papain substrate-pocket subsites S_4, S_3 and S_2, respectively [56]. This was further confirmed when the three-dimensional structure of cystatins was solved.

So far, only three 3D-structures of cystatins have been published: the crystallographic structures of N-truncated chicken cystatin C [57] and of a complex of human cystatin B (stefin B) with papain [58], as well as an NMR-structure of human cystatin A (stefin A) [59, 60]. From the structure of the complex between papain and stefin B, it is evident that the interactions between the enzyme and cystatins are formed by the amino-acid

Table 1. Dissociation constants K_i [nM] of enzyme–cystatin complexes

Cystatin	Cysteine protease				
	Papain	Cathepsin B	Cathepsin H	Cathepsin L	Cathepsin S
A	0.019[a]	8.2[a]	0.31[a]	1.3[a]	0.05[a]
B	0.12[a]	73[a]	0.58[a]	0.23[a]	0.07[a]
C	0.000011[b]	0.25[a]	0.28[a]	<0.005[a]	0.008[a]
D	1.2[a]	>1000[a]	8.5[a]	25[a]	0.24[a]
E	0.39[c]	32[c]	–	–	–
F	1.1[d]	>1000[d]	–	0.31[d]	–
S	108[a]	–	–	–	–
SA	0.32[a]	–	–	–	–
SN	0.016[a]	19[a]	–	–	–
H-kininogen	0.02[e]	400[f]	1.1[f]	0.109[f]	–
L-kininogen	0.015[a]	600[a]	0.72[a]	0.017[a]	–

[a]Ref. [52]; [b]calculated from association and dissociation constant velocities [53]; [c]ref. [38]; [d]ref. [40]; [e]ref. [54]; [f]ref. [55].

residues from the N-terminal segment (occupying S_n subsites of the enzyme) as well as by two additional fragments in β-hairpin loops: one in the middle and one in the C-terminal segment of the protein. These three cystatin regions, containing evolutionarily conserved amino-acid residues (Table 2), form a wedge-like structure, which interacts with the catalytic cleft of cysteine proteases [58]. It has been also proposed that the hydrophobic amino-acid residues from the first loop, as well as the tryptophan residue from the second loop, occupy the S_n' subsites of the enzyme. Kininogens, which have three cystatin-like domains, display also high affinity for papain and cathepsins

Table 2. Conserved amino-acid residues in binding segments of human cystatins[a]

Cystatin		N-terminus	I-loop	II-loop
A		MIP**G**G	**QVVAG**	
B		AcMMC**G**A	**QVVAG**	
C		RLV**G**G	**QIVAG**	**VPWQ**
D		TLA**G**G	**QIVAG**	**VPWE**
E		RMV**G**E	**QLVAG**	**VPWQ**
F		VKP**G**F	**QIVKG**	**VPWL**
S		IIP**G**G	**Q**TFG**G**	**VPWE**
SA		IIE**G**G	**QIVGG**	**VPWE**
SN		IIP**G**G	**QTVGG**	**VPWE**
H-kininogen	1-domain	*(QESQS)*	*(TVGSD)*	*(RSST)*
	2-domain	DCL**G**C	**Q**VV**AG**	(*DIQL*)
	3-domain	ICV**G**C	QVV**AG**	**VPW**E
L-kininogen	1-domain	*(QESQS)*	*(TVGSD)*	*(RSST)*
	2-domain	DCL**G**C	QVV**AG**	(*DIQL*)
	3-domain	ICV**G**C	QVV**AG**	**VPW**E

[a]Sequences in parenthesis correspond to the appropriate binding sequences of cystatins.

(Table 1). However, the first domain, which lacks the evolutionarily conserved N-terminal and β-hairpin-loop residues (Table 2, sequences in parenthesis), has no inhibitory activity against cysteine proteases.

HUMAN CYSTATIN C (hCC)

Human cystatin C (hCC; also called γ-trace, post-γ-globulin, gamma-CSF and post-gamma protein) was the first cystatin to be sequenced [42]. It is recognized as the most physiologically important extracellular human cystatin. Its primary structure consists of a single non-glycosylated polypeptide chain of 120 amino-acid residues (Fig. 3). The cysteine residues at positions 73 and

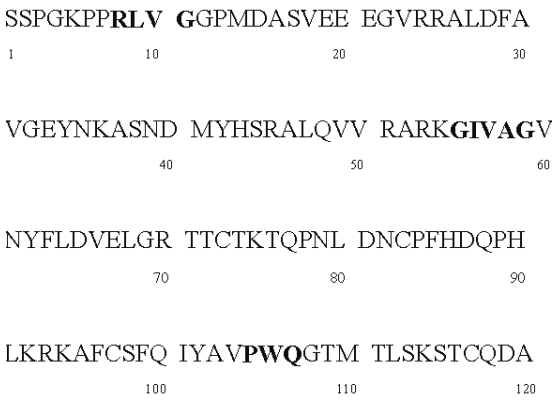

Figure 3. Primary structure of human cystatin C (hCC).

83 and those at positions 97 and 117 form two internal disulfide bridges. The nucleotide sequence of hCC has been determined and localized on chromosome 20 [61, 62]. Human cystatin C is present in all extracellular fluids. The highest concentration was found in seminal plasma (50 mg/L) [63], whereas normal blood plasma contains 0.8–2.5 mg/L of hCC [64]. The level of serum cystatin C is used now as an endogenous marker of renal function [64, 65]. hCC is an effective reversible inhibitor of cathepsins B, H, K, L and S [52]. The affinity of hCC for papain is too high to be measured by equilibrium methods. Therefore, the dissociation constant, $K_D = 1.1 \times 10^{-14}$ M (Table 1), for the hCC–papain complex was calculated from the association and dissociation rate constants [53].

Structure–activity relationship studies (SAR)

Cystatins contain three segments which are recognized as responsible for the interaction with cysteine proteases. These are the N-terminal fragment and the so-called first and second loops, which are arranged at one edge of the molecule and are believed to directly interact with the catalytic cleft of CPs. It has been shown in early SAR studies that truncation at the Gly^{11}–Gly^{12} peptide bond decreases the affinity of hCC for papain by three orders of magnitude [44, 66]. The importance of the N-terminal segment of hCC for its interaction with CPs was further confirmed by the studies of the rate of hydrolysis of appropriate synthetic peptides. Fragments comprising residues Gly^4–Glu^{21}, Arg^8–Asp^{15}, and Arg^8–Gly^{12} were all cleaved completely by papain at the Gly^{11}–Gly^{12} bond within less than 60 s, whereas the corresponding bond of the peptide comprising residues Gly^{11}–Asp^{15} was uncleaved even after 15-h incubation [56]. From these data, we postulated that the N-terminal fragment of hCC is involved in the inhibitor–enzyme interaction, and that the major contribution to the total affinity is through the binding of inhibitor residues Arg^8, Leu^9, and Val^{10} in the substrate subsites S_4, S_3 and S_2 of the enzyme [56]. This was further corroborated by SAR studies with hCC variants [66–68]. The side chain of Val^{10} has the most important contribution to the affinity of the N-terminal fragment of hCC for cathepsins. It was also shown that Leu^9 is the most discriminating residue for selective binding of hCC to cathepsins B, H, L, and S [68]. Exchange of the absolutely conserved Gly^{11} residue for other amino acids generally leads to a sharp decrease of the inhibitory potency [68], indicating that this residue may function as a hinge between the conformationally flexible N-terminal segment and the rest of the molecule [69, 70].

Structure–activity relationship for the remaining two binding segments of hCC (Gln^{55}–Gly^{59} and Pro^{105}–Trp^{106}) has been studied less extensively. Substitution of Trp^{106} by Gly decreases

the affinity for cathepsin B and H by approximately three orders of magnitude [68, 69]. The $Trp^{106} \rightarrow Gly^{106}$ substitution, when combined with a change in the N-terminal sequence of hCC, leads to a further sharp decrease of the inhibitory potency [66].

Leu68Gln mutation

One point mutation with glutamine residue substituting leucine at position 68 of hCC (Leu68Gln) is now recognized as a disease-causing disorder which leads to amyloid deposits in cerebral blood vessels [71–74]. This disorder known as hereditary cystatin C amyloid angiopathy (HCCAA) results in paralysis and development of dementia due to multiple strokes and death [74]. Indeed, it was shown that under various conditions the Leu68Gln mutant displays much higher tendency to dimerize and aggregate than wild-type hCC [75–77].

Dimerization, oligomerization

Early studies on thermal stability have shown that human cystatin C readily undergoes dimerization with complete loss of its inhibitory activity [77]. At a temperature above 80°C hCC aggregates, and consequently precipitates. Self-association of hCC was further evident when the protein was treated with various denaturing agents [76, 77]. NMR studies of human cystatin C have shown that it can form dimers through structural changes in its native fold [69, 77].

We have studied the structural changes of hCC occurring during both thermal and chemical denaturation processes. Chemical denaturation (with guanidine hydrochloride, Gdn·HCl) was examined by two spectroscopic methods: circular dichroism (CD) and tryptophan fluorescence [78]. To observe protein unfolding induced by heating, Fourier-transform infrared spectroscopy (FT-IR) was applied.

The obtained results indicate that unfolding of cystatin C caused by a denaturing agent is a complex process, characterized by two transition states (Fig. 4). The first one appeared in the concentration range of 0.5–1 M Gdn·HCl, the same as that interpreted by Ekiel & Abrahamson [77] as indicating the existence of dimeric cystatin C in their NMR studies. Thus, it can be concluded that the intermediate detected in our measurements is also a dimer.

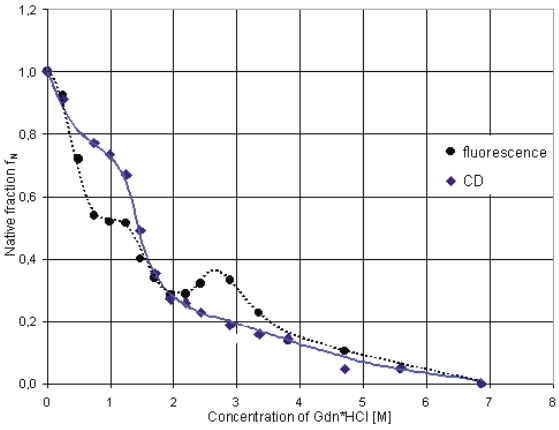

Figure 4. Conformational changes of cystatin C monitored by CD and tryptophan fluorescence analysis.

The fluorescence intensity was measured at 360 nm using an excitation wavelength of 295 nm. The ordinate shows the fraction of the native state calculated according to the equation $f_N = (x - x_D)/(x_N - x_D)$ where x is the value for spectroscopic parameter (ellipticity or fluorescence intensity) and x_N and x_D are the values for the native and denaturated states, respectively.

In the first transition state we did not detect any changes in the tertiary structure of cystatin C. Also very few changes were observed in the α-helix content. The only secondary structure motif exhibiting conformational changes after dimer formation, was the β-sheet. The most probable explanation of this fact is that β-strands participate directly in the formation of the dimeric molecule. After reverse conversion of the dimers into monomeric molecules, a dramatic loss of β-sheet content connected with the changes in the secondary and tertiary structure of cystatin C occurred. At a concentration of 2.5–3 M Gdn·HCl the second transition state, stabilized by partially recovered tertiary interactions, was detected (Fig. 4). However, it was only a temporary state preceding complete unfolding of cystatin C.

To study thermal denaturation of cystatin C, FT-IR spectroscopy was applied. The measure-

ments were performed for dry protein prepared by evaporation of a water solution. Oberg & Fink have reported [79] that solvent evaporation should not change the protein structure in solution. To confirm this statement, we carried out experiments at 35°C for dry cystatin C and cystatin C dissolved in water. The results reveal that at 35°C the protein structure in both cases is almost the same (Table 3). However, at higher temperatures no conformational changes in solid cystatin C could be detected. The dry protein retained its native state during the whole heating process.

Table 3. Percentage contents of different secondary structure motifs in native cystatin C

	β-sheet	α-helix	random coil	turn	3_{10}-helix, open loop, turn
Cystatin C in H_2O	41.7	14.8	16.5	24.8	2.2
Solid cystatin C	43.4	14.6	18.2	21.1	2.8

Structural studies of hCC

Our early molecular modeling studies on human cystatin C [80] have shown that the energy-optimized structure of hCC is very close to the crystallographic structure of chicken cystatin [57]. The results of fluorescence studies indicated that the Trp^{106} residue is fully exposed to solvent. We found that, apart from Trp^{106}, the main contribution to fluorescence comes from Tyr^{62} and Tyr^{42}. The remaining tyrosine residues (Tyr^{34} and Tyr^{102}) are efficiently quenched as a result of energy transfer to the Cys^{97}–Cys^{112} disulfide bridge (Tyr^{34}) and tryptophan (Tyr^{102}) [80].

Development of a second generation of more effective, specific cysteine protease peptide inhibitors would be greatly facilitated by the knowledge of three-dimensional structure of hCC. Similarly, such a model is necessary for the elucidation of the pathophysiological background of the cerebral hemorrhage produced by hCC, particularly its L68Q variant.

Crystallographic and NMR studies of chicken cystatin [57, 81, 82], cystatin B in complex with papain [58], cystatin A [60], and human cystatin D [83], have shown a similar overall structure, with three regions implicated for interactions with the target enzymes. Those regions include the N-terminal segment and two hairpin loops, L1 and L2. The general fold of protein inhibitors belonging to the cystatin family has been defined by the crystal structure of chicken cystatin [57]. Its canonical features include a long $\alpha 1$ helix running across a large, five-stranded antiparallel β sheet. The connectivity within the β sheet is as follows: (N)-β1-(α1)-β2-L1-β3-(AS)-β4-L2-β5-(C), where AS is a broad "appending structure", rather unrelated to the compact core of the remaining part of the molecule and positioned on the opposite end of the β sheet relative to the N-terminus and the two short loops L1 and L2. The latter three elements are aligned in a wedge-like fashion in the inhibitory motif of cystatins. Chicken cystatin shows 41% sequence identity and 62.5% homology to hCC but the crystal structure corresponds to an N-truncated variant [57]. On the other hand, the eleven N-terminal amino-acid residues of hCC are important for its very high-affinity binding to papain [52] (K_i 11 fM) and to other cysteine proteases [28]. It has been shown that specific cleavage, by leukocyte elastase, of the single N-terminal Val^{10}–Gly^{11} bond of hCC results in seriously compromised affinities for such target enzymes as cathepsin B, H, and L [84].

It is interesting to compare the topology of cystatin with that of the intensely sweet plant protein, monellin. The structural similarity has been noted before in spite of the low sequence identity [48, 49]. However, natural monellin consists of two protein chains: chain B, corresponding to helix $\alpha 1$ and strands $\beta 1$ and $\beta 2$ (in the order β1-α1-β2), and chain A, corresponding to the remaining, prominent part of the β sheet. The N-terminus of chain A and C-terminus of chain B are close in space and seem to be the product of proteolytic cleavage of a single-chain protein [48]

in a region that corresponds to cystatin loop L1 of the inhibitory "wedge". An artificial tethered B-A protein retains the taste and conformation of natural monellin [49].

There are two disulfide bonds in human cystatin C (Cys^{73}–Cys^{83}, Cys^{97}–Cys^{117}) and in all other proteins of family 2 and 3 cystatins [85]. Both are located within the β region of the chicken protein structure, in the C-terminal half of the molecule that would correspond to chain A in monellin. The conservation of these two S–S bridges in family 2 and 3 cystatins [45, 86] may be interpreted as implicating their requirement for stable protein fold. However, there are no disulfide bridges in family 1 cystatins or in monellin. In the structure of chicken cystatin there are two β-bulges, in strands $\beta 2$ (Arg^{46}) and $\beta 5$ (Leu^{111}), of the β sheet. They are preserved in the other structural models of cystatins, and also in monellin. The "appending helix" of chicken cystatin is disputable. It is only loosely connected with the molecular core and in the segment Cys^{71}–Lys^{91} is very poorly defined. In particular, the Lys^{73}–Leu^{78} fragment was weakly defined and tentatively placed, while in electron density the Asp^{85}–Lys^{91} peptide was not defined at all as it is completely disordered. In spite of that, the Asp^{77}–Asp^{85} fragment was modeled as helix $\alpha 2$. This α helix is not seen in the preliminary structure of human cystatin D [83] or in the structurally homologous monellin. Also, in the NMR studies of cystatins [69], no helical conformation has been found for this fragment either in chicken or human cystatin C. It appears that this fragment must be rather disordered in solution.

Crystallization of human cystatin C has been a challenge for a long time. Recently, formation of single crystals in several forms has been reported [87]. For the crystallization experiments, hCC was produced in its full-length form by recombinant techniques in *Escherichia coli* [88]. This full-length wild-type protein crystallized in two forms, tetragonal ($P4_12_12$ or $P4_32_12$) and cubic (I432). Low-temperature synchrotron data are available for both forms at the originally reported resolution of 3.0 and 3.1 Å, respectively [87]. The notorious poor quality and limited resolution of X-ray diffraction by full-length hCC crystals, in spite of their perfect and beautiful appearance, may be indicative of structural disorder (N-terminus, appending structure) and/or of lack of homogeneity resulting from uncontrolled protein aggregation (oligomerization) in the crystallization solutions and possibly also in the crystals. It should be stressed, however, that hCC used for growing the crystals represented pure monomeric protein obtained by gel filtration as the final isolation step. The Matthews volume [89] calculated for the two forms of full-length hCC is indicative of the presence of multiple copies of the protein in the asymmetric unit. The propensity of hCC to crystallize with multiple copies of the molecule in the asymmetric unit, in combination with the additional possibilities offered by the point symmetry elements of the unit cells, may be also indicative of the tendency of the protein to oligomerize. Such oligomerization might reflect the amyloid-forming property of Leu68Gln cystatin C, as earlier observations demonstrate that both wild type and Leu68Gln-substituted cystatin C are capable of forming dimers [69, 74, 77]. In the tetragonal form, as many as seven independent molecules could be present. The cubic unit cell is likely to contain two asymmetric copies (V_m 2.16 $Å^3$/Da), but one molecule and high solvent content (72%) is also possible. To facilitate the solution of the crystal structure of hCC, the full-length protein was also produced in selenomethionyl form [87]. Electrospray mass spectrometry of the selenomethionyl protein confirmed that the three Met residues in the hCC sequence were fully substituted by Se-Met. A successful Met→Se-Met substitution was additionally confirmed by analysis of the amino-acid composition of the Se-Met protein after acidic hydrolysis. The selenomethionyl protein crystallized in the cubic form and X-ray absorption spectra confirmed a significant content of selenium in the crystals. Unfortunately, due to weak diffraction, only multiwavelength anomalous diffraction (MAD) data at 4.5 Å resolution could be measured for those crystals at the selenium absorption edge.

Very recently, a new low-temperature data set was obtained for the cubic form of native full-length hCC using synchrotron radiation (R. Janowski, unpublished). This data set extends to 2.5 Å resolution and is currently being used for

the determination of the structure of hCC. In addition to the experiments involving full-length human cystatin C, preliminary crystallographic studies have also been reported for its N-terminally truncated variant [87]. hCC devoid of ten N-terminal residues was obtained by incubation of recombinant wild type human cystatin C with leukocyte elastase and isolated as described by Abrahamson *et al.* [84]. The protein could be crystallized in tetragonal form yielding crystals that are very stable in the X-ray beam. Measurement of diffraction data extending to 2.7 Å has been reported at room temperature, using conventional Cu $K\alpha$ radiation [87]. Also in the case of N-truncated hCC, the asymmetric unit can be expected to contain numerous (up to eleven) independent copies of the protein. Very recently, a new diffraction data set extending to 2.1 Å resolution has been measured at low temperature using synchrotron radiation (R. Janowski, unpublished).

LOW-MOLECULAR-MASS INHIBITORS OF CYSTEINE PROTEASES RELATED TO THE STRUCTURE OF THE BINDING CENTER OF CYSTATINS

Peptidyl-diazomethyl ketones

Soon after the discovery that the N-terminal fragment: Arg^8-Leu^9-Val^{10}-Gly^{11} of human cystatin C interacts with the S_n subsites of cysteine proteases [53, 56], a series of peptidyl-diazomethyl ketones based on the structure of this segment was synthesized. Preliminary results showed that both Boc-Val-Gly-CHN_2 (Boc-VG-DAM) and Z-Leu-Val-Gly-CHN_2 (Z-LVG-DAM) inhibit papain, cathepsin B and streptococcal proteinase [56]. The latter compound was tested for *in vitro* and *in vivo* antibacterial activity against a large number of bacterial strains of different species [90]. Mice injected with lethal doses of group A streptococci were cured by a single injection of 0.2 mg of Z-LVG-DAM. Detailed structure-activity studies showed that the shortest among diazomethyl ketones, Z-Gly-CHN_2, does not inhibit cysteine proteases (Table 4). On the other hand, extention of the -Leu-Val-Gly- sequence by an Arg residue in Z-RLVG-DAM gave the most potent inhibitor of papain and cathepsin B, with apparent second order rate constants (k_{+2}') of the same order of magnitude as those determined for E-64, which is used as standard in the inhibitory bio-assays of cysteine proteases [91]. Addition of the next Pro^7 residue in Z-PRLVG-DAM decreased the activity. Peptidyl-diazomethyl ketones with a free N-terminal amino group displayed a lower inhibitory potency. None of the peptidyl-diazomethyl ketones designed after the N-terminal sequence of various cystatins had an inhibitory activity higher than that of hCC itself [91, 92]. These peptidyl-diazomethyl ketone inhibitors have been found to be very fast and irreversible inhibitors of cysteine proteases. It should be noted that the reactivity of the diazomethyl ketone group with thiols is generally very low [93]. Modified neglect of diatomic overlap (MNDO) studies of the mechanism of inhibition of cysteine proteases by diazomethyl ketones showed that the reaction is irreversible and leads to an α-thioketone derivative of the Cys^{25} residue of papain [94]. Recently, we have shown that Z-RLVG-DAM inhibits bone resorption *in vitro* by a mechanism that seems primarily due to inhibition of bone matrix degradation *via* cysteine proteases [95].

Oxirane-type inhibitors

E-64 [(2*S*,3*S*)-*trans*-epoxysuccinyl-L-leucyl-agmatine] isolated from cultures of *Aspergillus japonicus* is a very strong and irreversible inhibitor of cysteine proteases [96, 97]. The first oxirane-containing inhibitor, based on the structure of the N-terminal segment of hCC designed by us, Z-Leu-Val-$NHCH_2$-CH(O)CH-CH_2COOH, displayed only weak reversible inhibition [56]. Therefore, taking into account the structure of E-64 and its analogs, as well as our modeling studies, we have designed several new compounds with more hydrophobic C-termini. Most of these compounds displayed quite good inhibitory activities towards papain and cathepsin B. However, the most striking result came from two oxirane-type compounds: Z-Arg-Leu-ValΨ[CH_2NH]CO-

$CH(O)CH\text{-}C_6H_5$ (Table 4, compound **16**) and Z-Arg-Leu-ValΨ[CH_2NH]CO-$CH(O)CH\text{-}COC_6H_5$ (Table 4, compound **17**). Compound **17**, with a stronger electron-withdrawing benzoyl group at the C-terminus, was found to be a good irreversible inhibitor of papain and cathepsin B, whereas compound **16**, with the phenyl ring attached directly to the oxirane moiety, had no inhibitory potency towards cysteine proteases. This discrepancy prompted us to undertake more detailed

Table 4. Comparison of the inhibition rate constants ($M^{-1}s^{-1}$) for inhibitors of cysteine proteases containing N-terminal binding segments of cystatins

No.	Compound	Cysteine protease			*Streptococcus gr. A* killing effect
		Papain	Cathepsin B (bovine)	Other cathepsins	
Peptidyl-diazomethyl ketones		k'_{+2} [$M^{-1}s^{-1}$]			
1.	Z-Gly-DAM[a]	$<10^2$	inactive		
2.	Boc-Val-Gly-DAM[a]	1.66×10^4	1.21×0^3		
3.	Z-Leu-Val-Gly-DAM[a]	4.48×10^5	2.4×10^4	3.57×10^4 Cath. L 5.22×10^4 Cath. S $<10^2$ Cath. H	+++
4.	H-Leu-Val-Gly-DAM[a]	2.26×10^4	1.5×10^3		
5.	Z-Arg-Leu-Val-Gly-DAM[a]	5.59×10^5	3.96×10^4		inactive
6.	H-Arg-Leu-Val-Gly-DAM[b]	6.25×10^4			
7.	Z-Pro-Leu-Val-Gly-DAM[b]	4.22×10^5			
8.	Z-Ile-Val-Gly-DAM[b]	2.97×10^5	3.63×10^3		++
9.	Z-Leu-Leu-Gly-DAM[b]	1.56×10^4			++
10.	Z-Arg-Leu-Leu-Gly-DAM[b]	5.47×10^4	1.94×10^4		inactive
11.	Z-Leu-Ala-Gly-DAM[b]	1.09×10^4	1.98×10^3	9.3×10^2 Cath. L 1.82×10^4 Cath. S $<10^2$ Cath. H	++
12.	Z-Ile-Pro-Gly-DAM[b]	3.52×10^3	inactive		inactive
13.	Z-Arg-Ile-Ile-Pro-Gly-DAM[b]	1.47×10^3	inactive		inactive
14.	Z-Arg-Thr-Leu-Ala-DAM[b]	2.62×10^4			inactive
15.	Z-Phe-Leu-Gly-DAM[b]	1.16×10^4	3.50×10^3		++
	E-64[a]	8.13×10^5	5.35×10^4		
Oxiranes		k'_{+2} [$M^{-1}s^{-1}$]			
16.	Z-Arg-Leu-ValΨ[CH_2NH]CO-$CH(O)CH\text{-}C_6H_5$[b]	inactive	inactive		
17.	Z-Arg-Leu-ValΨ[CH_2NH]CO-$CH(O)CH\text{-}CO\text{-}C_6H_5$[b]	1.73×10^4	8.0×10^3		
18.	Z-Arg-Leu-ValΨ[CH_2NH]CO-$CH(O)CH\text{-}COOCH_3$[b]	2.67×10^4			
19.	dArg-Leu-ValΨ[CH_2NH]CO-$CH(O)CH\text{-}CO\text{-}C_6H_5$[b]	3.06×10^3	5.43×10^3		
20.	C_6H_5-CO-CH(O)CH-CO-Val-Leu-$NH(CH_2)_4NHC(NH)NH_2$[b]	K_i= 16.3 nM	K_i = 0.256 nM		

[a] [91]; [b] this paper.

structural studies using molecular modeling and crystallographic methods.

Other inhibitors

Apart from diazomethyl ketone- and oxirane-type inhibitors, we have designed several other compounds containing the cystatin binding motif, as well as reactive groups for the thiol function of cysteine proteases [56, 98]. Good inhibitory potency was found for compounds containing an activated olefinic double bond and compounds with a C-terminal aldehyde group or chloro- and bromomethyl ketone groups. It was interesting to find that most of them displayed antibacterial activity against seventeen clinically important bacterial species tested [98]. It should be mentioned that many cyclic peptides based on the N-terminal sequence of cystatin C also displayed antibacterial properties. Recently, we have designed and synthesized several azapeptides based on the binding sequence of cystatins, and some of them were found to be very selective inhibitors of different cathepsins (E. Wieczerzak, unpublished).

Crystallographic studies of papain–inhibitor complexes

Single crystals of the covalent complex papain—Z-Arg-Leu-ValΨ[CH_2NH]-CO-CH(O)CH-COC_6H_5 (Table 4, compound **17**) were grown by the vapor diffusion method at room temperature in hanging drops using a modification of the procedure described for the complex papain—E-64c [99]. Detailed crystallization conditions and the procedure for data collection at room temperature using freshly grown crystals (data set I — resolution 1.9 Å) were described previously [100]. Another, low-temperature data set was collected about 10 months later using synchrotron radiation (resolution 1.65 Å). The crystals used in those studies correspond to the historically first crystal form of papain, form A, crystallized by Drenth & Jansonius [101], for which no crystal structure has yet been reported.

Even preliminary difference electron density maps calculated using the room-temperature data (data set I) clearly showed the inhibitor, which is covalently linked to the active-site Cys^{25} of the enzyme. However, the maps calculated using the low-temperature data (data set II) clearly revealed only a short stem of electron density near the active-site Cys^{25}. An analysis of the shape of this electron density and of potential hydrogen bonds strongly suggests that in these aged crystals of the complex, the inhibitor that was originally attached to the sulfhydryl group of Cys^{25} has been replaced by a covalent hydroxyethyl substituent.

The overall structure (room- and low-temperature models) of the enzyme is similar to other papain structures deposited in the PDB, and the r.m.s. deviation for Cα atoms of these two models is 0.24 Å.

The inhibitor moiety in the room-temperature structure extends along the S_n subsites of the enzyme (Fig. 5) and is stabilized in the active-site groove by a series of hydrogen bonds and hydrophobic interactions. The inhibitor forms hydrogen bonds with Gly^{66}, Asp^{158}, and Gln^{19} as well as with two solvent molecules. Similar contacts were also observed in the 2.1 Å resolution structure of a complex between papain and E-64c [99]. The hydrophobic interactions with the S_2 subsite characteristic for chloromethylketone inhibitors were not observed. The distances between the side chains of Val^{133} and Val^{157} (defining the enzyme's S_2 subsite) and the atoms of the Val residue of the inhibitor, are longer than 6.0 Å.

As a step towards understanding the specificity of peptidic, covalent, irreversible inhibitors of papain, two peptidyl-diazomethyl ketone-type inhibitors: Z-Arg-Leu-Val-Gly-DAM and Z-Leu-Phe-Gly-DAM (Table 4), with valine and phenylalanine residues in the P_2 site, respectively, were synthesized and reacted with the active site of papain. The complex between papain and the Z-Arg-Leu-Val-Gly-DAM has been characterized crystallographically (space group $P2_1$, 1.78 Å resolution, R = 0.168). The side chain of Val from the Z-Arg-Leu-Val-Gly-DAM inhibitor molecule is rather far from the hydrophobic S_2 pocket, the closest distances in this region being above 4.6 Å. Electron density is clearly visible for the entire inhibitor moiety with the exception of the benzyloxycarbonyl (Z) group. The structure, therefore, demonstrates

again no specific association between the S_2 pocket and the inhibitor's P_2 site, analogously to the situation observed in the crystal structure of the Z-Arg-Leu-ValΨ[CH_2NH]CO-CH(O)CH-CO-C_6H_5 complex [100], and in molecular dynamics simulations [102]. This persistent lack of P_2–S_2 interactions in Z-Arg-Leu-Val-type inhibitors is in contrast to the early findings by Drenth *et al.* [103] that P_2–S_2 complementarity is essential for productive inhibition and for enzyme specificity. This evidence seems to indicate that, while it might be important for efficient and precise docking of the inhibitor in the active site, the S_2 pocket does not play any significant role in the association between the inhibitor and the enzyme once a covalent bond has been formed.

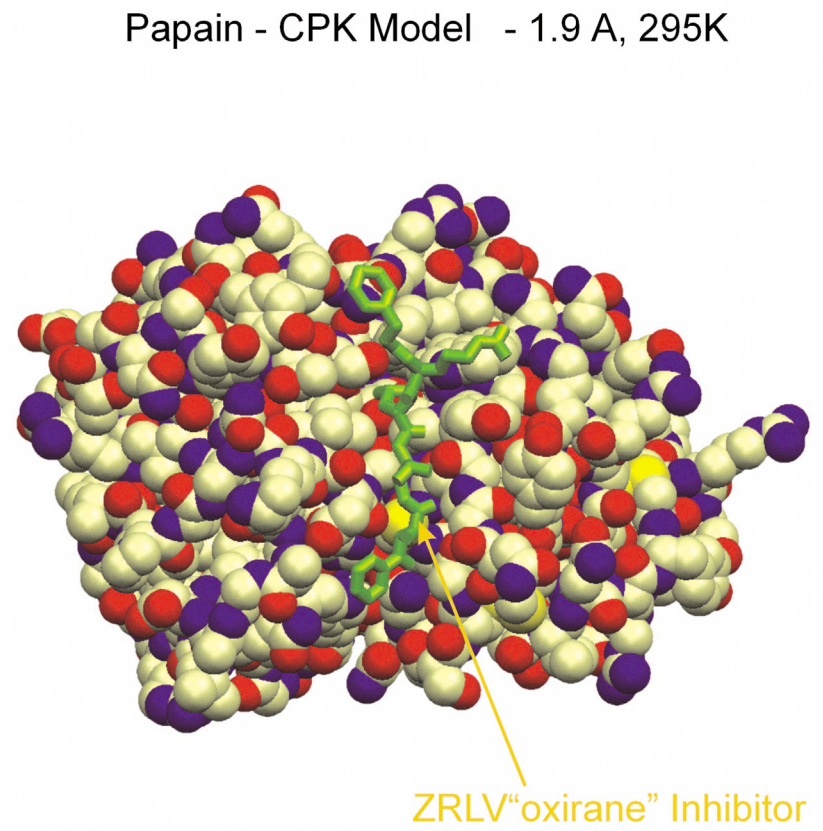

Figure 5. Crystal structure of the papain–oxirane inhibitor (17) complex (M. Kozak, unpublished).

The enzyme is shown as space-filling model viewed into the catalytic cleft from the outside. The inhibitor (green) is seen in the cleft in its fully extended conformation, with the Z group at the top and the "oxirane" moiety, now opened and covalently linked to the enzyme's Cys^{25} Sγ atom (yellow), at the bottom.

Two polymorphs of a complex between papain and the Z-Leu-Phe-Gly-DAM inhibitor have been crystallized. Diffraction data for crystal form I (space group $P2_1$) were collected to 2.0 Å resolution, and for crystal form II (space group $P2_12_12_1$) to 1.63 Å. Both structures were solved by molecular replacement using the 1ppn.pdb [104] model of papain as a probe. The final R factors are 0.106 and 0.172, respectively. In both crystal forms the inhibitor is bound to the Cys^{25} residue of the papain with a covalent bond formed between the methylene group (DAM) of the inhibitor and the thiol group of the enzyme. The phenylalanyl side chain is locked by hydrophobic interactions (3.5–3.8 Å) with residues Val^{133} and Val^{157} of the S_2 pocket. The orientation of the phenyl ring is similar to that observed in the chloromethyl ketone complexes studied by Drenth *et al.* (Protein Data Bank codes: 1pad, 5pad, 6pad) [103]. The inhibitor is stabilized by additional hydrogen bonds between its main chain and the residues forming the catalytic cleft of the enzyme (highly conserved hydrogen bonds with Gln^{19} and Gly^{66}). Electron density is very clear for the covalent bond connecting the inhibitor and the enzyme as well as for the side chain of the phenylalanyl residue. The N-terminal part of the inhibitor, the benzyloxycarbonyl group (Z), has no visible electron density in either of the crystal forms.

Based on the above structures and the structures of other papain–inhibitor complexes one can conclude that, in covalent papain–inhibitor complexes, hydrophobicity of the P_2 residue is not

sufficient for productive binding of the inhibitor in the S_2 pocket and that its bulkiness is equally important. This does not preclude, however, that even a smaller residue, like valine, may be effective during recognition and docking prior to the formation of the covalent link.

Molecular modeling

The initial models for papain–Z-Arg-Leu-ValΨ[CH_2NH]-CH(O)CH-COC_6H_5 (Z-R-L-Nnv-Oxi–papain; papain–InhA) and papain–Z-Arg-Leu-Val-Gly-CHN_2 (Z-RLVG-CH_2–papain; papain–InhB) complexes were prepared in the Sybyl program [105] based on papain Protein Data Bank files 1pe6 and 1pad, respectively [106], in combination with the respective ligands modeled as described elsewhere ([102] and P. Drabik, unpublished). The complexes were subsequently analysed using the AMBER program [107]. New residues, absent in the original AMBER force field were parameterized according to standard procedures [108].

The initial structures of the papain–ligand complexes were subjected to constrained simulated annealing [102] enabling the simulation at very high, physically unrealistic temperature. The additional kinetic energy enhanced the ability of the system to explore the energy surface and to avoid getting stuck in energetically unfavourable local energy minima. Afterwards, the systems were subjected to 230 ps of unconstrained molecular dynamics at 300 K.

Time-averaged residue-based deviations as a function of residue number for all molecular dynamics runs indicated (Fig. 6) changes up to 5 Å for some residues (C-terminus), but the overall C mobilities oscillated about 1 Å for InhA and 2 Å for InhB. The average structures displayed similarly significant changes in some flexible loops of the enzyme. It should be stressed that the protein structures reproduced very well the mobility pattern typical of the whole molecule as represented by the atomic displacement parameters (temperature factors) in the crystal structure of the papain–E-64c complex [100]. This result validated the use of the AMBER 5.0 force field as a suitable tool for scanning the conformational space of both ligands in the catalytic cleft of papain.

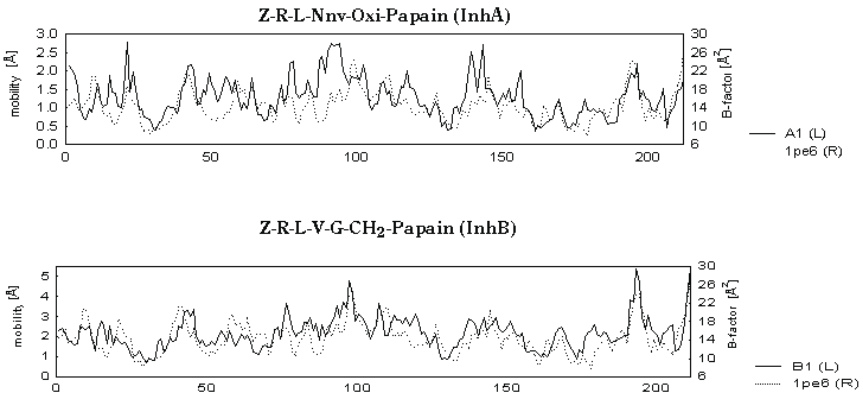

Figure 6. Time-averaged residue-based deviations (mobilities) along the papain sequence during molecular dynamics (MD) runs.

Symbols A1 and B1 correspond to InhA and InhB MD runs, respectively. The distribution of both quantities along the sequence is evident. It can be seen that some loops of the papain L lobe undergo high fluctuations during molecular dynamics runs.

Detailed atomic-level analysis of the mobilities of the inhibitor backbones reveals that the scatter of the relaxed positions of the residues increases steeply towards the N-terminus of the inhibitors (Fig. 7). Thus, the catalytic pocket S_3, as defined by the pioneering studies of Schechter & Berger

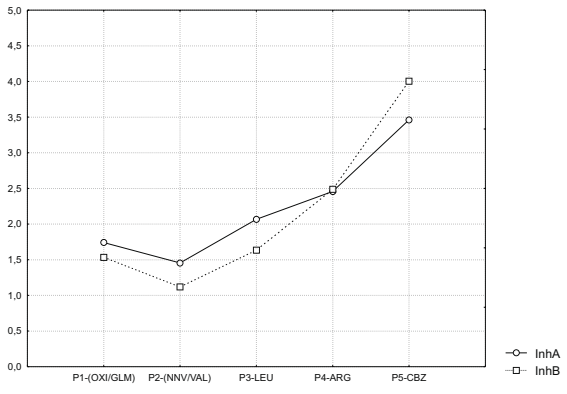

Figure 7. Mobilities of inhibitor residues during molecular dynamics simulations.

[6], appears rather elusive in view of the inhibitor flexibility evident from the molecular dynamics simulations and from the experimentally determined structures of papain–inhibitor complexes [58, 99, 103, 109, 110]. The location and definition of the substrate binding site S_4 is even more questionable.

REFERENCES

1. Barrett, A.J. (1994) Classification of peptidases. *Methods Enzymol.* **244**, 1–15.

2. Hartley, B.S. (1960) Proteolytic enzymes. *Annu. Rev. Biochem.* **9**, 45–72.

3. Otto, H.-H. & Schirmeister, T. (1997) Cysteine proteases and their inhibitors. *Chem. Rev.* **97**, 133–171.

4. Rawlings, N.D. & Barrett, A.J. (1999) MEROPS: The peptidase database. *Nucleic Acids Res.* **27**, 325–331.

5. McGrath, M.E. (1999) The lysosomal cysteine proteases. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 181–204.

6. Schechter, I. & Berger, A. (1967) On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162.

7. Turk, D., Gunčar, G., Podobnik, M. & Turk, B. (1998) Revised definition of substrate sites of papain-like cysteine proteases. *Biol. Chem.* **379**, 137–147.

8. Berti, P.J. & Storer, A.C. (1995) Alignment/phylogeny of the papain superfamily of cysteine proteases. *J. Mol. Biol.* **246**, 273–283.

9. Turk, B., Turk, V. & Turk, D. (1997) Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol. Chem.* **378**, 141–150.

10. Storer, A.C. & Menard, R. (1994) Catalytic mechanism in papain family of cysteine peptidases. *Methods Enzymol.* **244**, 486–500.

11. Kirschke, H., Barrett, A.J. & Rawlings, N.D. (1995) Proteinases 1: Lysosomal cysteine proteinases; in *Proteine Profile* 2 (Sheterline, P., ed.) pp. 1587–1643, Oxford University Press.

12. Afonso, S., Romagnano, L. & Babiarz, B. (1997) The expression and function of cystatin C and cathepsin B and cathepsin L during mouse embryo implantation and placentation. *Development* **124**, 3415–3425.

13. Henskens, M.C., Veerman, E.C.I. & Amerongen, A.V.N. (1996) Cystatins in health and disease. *Biol. Chem. Hoppe-Seyler* **377**, 71–86.

14. Grubb, A. (2000) Cystatin C – properties and use as diagnostic marker. *Adv. Clin. Chem.* **35**, 63–99.

15. Warwas, M. & Haczyńska, H. (1998) Rola cystatyn w procesie nowotworowym i jego diagnostyce. *Post. Hig. Med. Dośw.* **52**, 515–526 (in Polish).

16. Machleidt, W., Borchart, U., Fritz, H., Brzin, J., Ritonja, A. & Turk, V. (1983) Protein inhibitors of cysteine proteinases: II. Primary structure of stefin, a cytosolic protein inhibitor of cysteine proteinases from human polymorphonuclear granulocytes. *Hoppe-Seyler's Physiol. Chem.* **364**, 1481–1486.

17. Green, G.D.J., Kembhavi, A.A., Davies, M.E. & Barrett, A.J. (1984) Cystatin-like proteinase inhibitors from human liver. *Biochem. J.* **218**, 939–946.

18. Ritonja, A., Machleidt, W. & Barrett, A.J. (1985) Amino acid sequence of the intracellular cysteine proteinase inhibitor cystatin B from human liver. *Biochem. Biophys. Res. Commun.* **131**, 1187–1192.

19. Takeda, A., Kobayashi, S. & Samejima, T. (1983) Isolation and characterization of two thiol proteinase inhibitors of low molecular weight from newborn rat epidermis. *J. Biochem.* **94**, 811–820.

20. Turk, B., Kri[aj, I. & Turk, V. (1992) Isolation and characterization of bovine stefin B. *Biol. Chem. Hoppe-Seyler* **373**, 441–446.

21. Turk, B., Ritonja, A., Björk, I., Stoka, V., Dolenc, I. & Turk, V. (1995) Identification of bovine stefin A, a novel inhibitor of cysteine proteinases. *FEBS Lett.* **360**, 101–105.

22. Lenarčič, B., Ritonja, A., Dolenc, I., Stoka, V., Berbic, S., Pungercar, J., Strukelj, P. & Turk, V. (1993) Pig leukocyte cysteine proteinase inhibitor (PLCPI), a new member of stefin family. *FEBS Lett.* **336**, 289–292.

23. Kondo, H., Ijjri, S., Abe, K., Maeda, H. & Arai, S. (1992) Inhibitory effect of oryzacystatins and a truncation mutant on the replication of poliovirus in infected Vero cells. *FEBS Lett.* **299**, 48–50.

24. Lenarčič, B., Kri[aj, I., Zunec, P. & Turk, V. (1996) Differences in specificity for the interactions of stefins A, B and D with cysteine proteinases. *FEBS Lett.* **395**, 113–118.

25. Rawlings, N.D. & Barrett, A.J. (1990) Evolution of proteins of the cystatin superfamily. *J. Mol. Evol.* **30**, 60–71.

26. Lenarčič, B., Ritonja, A., Sali, A., Kotnik, M., Turk, V. & Machleidt, W. (1986) Properties and structure of human spleen stefin B — a low molecular weight protein inhibitor of cysteine proteinases; in *Cysteine Proteinases and their Inhibitors* (Turk, V., ed.) pp. 473–487, Walter de Gruyter.

27. Katunuma, N. & Kominami, E. (1986) Distribution and localization of lysosomal cysteine proteinases and cystatins; in *Cysteine Proteinases and their Inhibitors* (Turk, V., ed.) pp. 219–227, Walter de Gruyter.

28. Abrahamson, M. (1993) Cystatins — protein inhibitors of papain-like cysteine proteinases. *Ciencia e Cultura* **45**, 229–304.

29. Hsieh, W.-T., Fong, D., Sloane, B.F., Golembieski, W. & Smith, D.I. (1991) Mapping of the gene for human cysteine proteinase inhibitor stefin A, STFI, to chromosome 3cen-q21. *Genomics* **9**, 207–209.

30. Pennacchio, L.A., Lahesjoki, 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) Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1). *Science* **271**, 1731–1734.

31. Fossum, K. & Whitaker, J.R. (1968) Ficin and papain inhibitor from chicken egg white. *Arch. Biochem. Biophys.* **125**, 367–375.

32. Sen, L.C. & Whitaker, J.R. (1973) Some properties of a ficin-papain inhibitor from avian egg white. *Arch. Biochem. Biophys.* **158**, 623–632.

33. Saitoh, E., Sabatini, L., Eddy, R.L., Shows, T.B., Azen, S., Isemura, S. & Sanada, K. (1989) The human cystatin C gene (CST3) is a member of the cystatin gene family which is localized on chromosome 20. *Biochem. Biophys. Res. Commun.* **162**, 1324–1331.

34. Freije, J.P., Pendas, A.M., Velasco, G., Roca, A., Abrahamson, M. & Lopez-Otin, C. (1993) Localization of the human cystatin D gene (CST5) to human chromosome 20p11.21 by *in situ* hybridization. *Cytogen. Cell. Genet.* **62**, 29–31.

35. Isemura, S., Saitoh, E., Sanada, K. & Minekata, K. (1991) Identification of full-sized forms of salivary (S-type) cystatins (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of cystatin S) in human whole saliva and determination of phosphorylation sites of cystatin S. *J. Biochem.* (Tokyo) **110**, 648–654.

36. Balbin, B., Freije, J.P., Abrahamson, M., Velasco, G., Grubb, A. & Lopez-Otin, C. (1993) A sequence variation in the human cystatin D gene resulting in an amino acid (Cys/Arg) polymorphism at the protein level. *Hum. Genet.* **90**, 668–669.

37. Freije, J.P., Balbin, M., Abrahamson, M., Velasco, G., Dalbøge, M., Grubb, A. & Lopez-Otin, C. (1993) Human cystatin D. cDNA cloning, characterization of the *Escherichia coli* expressed inhibitor, and identification of the native protein in saliva. *J. Biol. Chem.* **268**, 15737–15744.

38. Ni, J., Abrahamson, M., Zhang, M., Alvarez-Fernandez, M., 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) Cystatin E is a novel human cysteine proteinase inhibitor with structural resemblance to family 2 cystatins. *J. Biol. Chem.* **272**, 10853–10858.

39. Sotiropoulou, G., Anisonowicz, A. & Sager, R. (1997) Identification, cloning, and characterization of cystatin M, a novel cysteine proteinase inhibitor, down-regulated in breast cancer. *J. Biol. Chem.* **272**, 903–910.

40. Ni, J., Fernandez, M.A., Danielsson, L., Chillakuru, R.A., Zhong, J., Grubb, A., Su, J., Gentz, R. & Abrahamson, M. (1998) Cystatin F is a glycosylated human low molecular weight cysteine proteinase inhibitor. *J. Biol. Chem.* **273**, 24797–24804.

41. Halfon, S., Ford, J., Foster, J., Dowling, L., Lucian, L., Sterling, M., Xu, Y., Weiss, M., Ikeda, M., Liggett, D., Helm, A., Caux, C., Lebecques, S., Hannum, C., Menon, S., Mc Clanahan, T., Gorman, D. & Zurawski, G. (1998) Leukocystatin, a new class II cystatin expressed selectively by hematopoietic cells. *J. Biol. Chem.* **273**, 16400–16408.

42. Grubb, A. & Lofberg, H. (1982) Human γ-trace, a basic microprotein: Amino acid sequence and pres-

ence in the adenophysis. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3024–3027.

43. Abrahamson, M., Grubb, A., Olafsson, I. & Lundwall, A. (1987) Molecular cloning and sequence analysis of cDNA coding for the precursor of the human cysteine proteinase inhibitor cystatin C. *FEBS Lett.* **216**, 229–233.

44. Abrahamson, M., Ritonja, A., Brown, M.A., Grubb, A., Machleidt, W. & Barrett, A.J. (1987) Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. *J. Biol. Chem.* **262**, 9688–9694.

45. Abrahamson, M., Barrett, A.J., Salvessen, G. & Grubb, A. (1986) Isolation of six cysteine proteinase inhibitors from human urine. Their physicochemical and enzyme kinetic properties and concentrations in biological fluids. *J. Biol. Chem.* **261**, 11282–11289.

46. DeLa Cadena, R.A. & Colman, R.N. (1991) Structure and functions of human kininogens. *Trends Pharmacol. Sci.* **12**, 272–275.

47. Brown, W.M. & Dziegielewska, K.M. (1997) Friends and relations of the cystatin superfamily – New members and their evolution. *Protein Sci.* **6**, 5–12.

48. Murzin, A. (1993) Sweet-tasting protein monellin is related to the cystatin family of thiol proteinase inhibitors. *J. Mol. Biol.* **230**, 680–694.

49. Somoza, J.R., Jiang, F., Tong, L., Kang, C.-H., Cho, J.M. & Kim, S.-H. (1993) The crystal structures of a potently sweet protein. Natural monellin at 2.75 Å resolution and a single-chain monellin at 1.7 Å resolution. *J. Mol. Biol.* **234**, 390–404.

50. Bode, W., Engh, R., Musil, D., Laber, B., Stubbs, M., Huber, R. & Turk, V. (1990) Mechanism of interaction of cysteine proteinases and their protein inhibitors as compared to the serine proteinase–inhibitor interaction. *Biol. Chem. Hoppe-Seyler* **371**, 111–118.

51. Björk, I. & Ylinenjärvi, K. (1989) Interaction of chicken cystatin with inactivated papains. *Biochem. J.* **260**, 61–68.

52. Abrahamson, M. (1994) Cystatins. *Methods Enzymol.* **244**, 685–700.

53. Lindahl, P., Abrahamson, M. & Björk, I. (1992) Interaction of recombinant human cystatin C with cysteine proteinases papain and actinidin. *Biochem. J.* **28**, 49–55.

54. Müller-Esterl, W., Fritz, H., Machleidt, W., Ritonja, A., Brzin, J., Kotnik, M., Turk, V., Kellermann, J. & Lottspeich, F. (1985) Human plasma kininogens are identical with α-cysteine proteinase inhibitors. *FEBS Lett.* **182**, 310–314.

55. Barrett, A.J., Rawlings, N.D., Davies, M.E., Machleidt, W., Salvesen, G. & Turk, V. (1986) Cysteine proteinase inhibitors of the cystatin superfamily; in *Proteinase Inhibitors* (Barrett, A.J. & Salvesen, G., eds.) pp. 515–569, Springer-Verlag.

56. Grubb, A., Abrahamson, M., Olafsson, I., Trojnar, J., Kasprzykowska, R., Kasprzykowski, F. & Grzonka, Z. (1990) Synthesis of cysteine proteinase inhibitors structurally based on the proteinase interacting N-terminal region of human cystatin C. *Biol. Chem. Hoppe-Seyler* **371** (Suppl.), 137–144.

57. Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshnikov, A., Brzin, J., Kos, J. & Turk, V. (1988) The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.* **7**, 2593–2599.

58. Stubbs, M.T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarčič, B. & Turk, V. (1990) The refined 2.4 Å X-ray crystal structure of the recombinant human stefin B in complex with the cysteine proteinase papain: A novel type of proteinase inhibitor interaction. *EMBO J.* **9**, 1939–1947.

59. Martin, J.R., Jerala, R., Kroon-Zitko, L., [erovnik, E., Turk, V. & Waltho, J.P. (1994) Structural characterization of human stefin A in solution and implications for binding to cysteine proteinases. *Eur. J. Biochem.* **225**, 1181–1194.

60. Martin, J.R., Craven, C.J., Jerala, R., Kroon-Zitko, L., [erovnik, E., Turk, V. & Waltho, J.P. (1995) The three-dimensional solution structure of human stefin A. *J. Mol. Biol.* **246**, 331–343.

61. Abrahamson, M., Olafsson, I., Palsdottir, A. Ulvsbäck, M., Lundwall, Å., Jensson, O. & Grubb, A. (1990) Structure and expression of the human cystatin C gene. *Biochem. J.* **268**, 287–294.

62. Schnittger, S., Gopal Rao, V.V.N., Abrahamson, M. & Hansmann, L. (1993) Cystatin C (CST3), the candidate gene for hereditary cystatin C amyloid angiopathy (HCCAA), and other members of the cystatin gene family are clustered on chromosome 20p11.2. *Genomics*, **16**, 50–55.

63. Grubb, A., Weiber, H. & Löfberg, H. (1983) The γ-trace concentration of normal human seminal plasma is thirty-six times that of normal human blood plasma. *Scand. J. Clin. Lab.* **43**, 421–425.

64. Grubb, A. (1992) Diagnostic value of analysis of cystatin C and protein HC in biological fluids. *Clin. Nephrol.* **38** (Suppl. 1), S20–S27.

65. Randers, E. & Erlandsen, E.J. (1999) Serum cystatin C as an endogenous marker of the renal function — a review. *Clin. Chem. Lab. Med.* **37**, 389–395.

66. Hall, A., Håkansson, K., Mason, R.W., Grubb, A. & Abrahamson, M. (1995) Structural basis for the biological specificity of cystatin C. Identification of leucine 9 in the N-terminal binding region as a selectivity-conferring residue in the inhibition of mammalian cysteine peptidases. *J. Biol. Chem.* **270**, 5115–5121.

67. Mason, R.W., Sol-Church, K. & Abrahamson, M. (1998) Amino acid substitutions in the N-terminal segment of cystatin C create selective protein inhibitors of lysosomal cysteine proteinases. *Biochem. J.* **330**, 833–838.

68. Hall, A., Dalbøge, H., Grubb, A. & Abrahamson, M. (1993) Importance of the evolutionary conserved glycine residue in the N-terminal region of human cystatin C. *Biochem. J.* **291**, 123–129.

69. 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) NMR structural studies of human cystatin C dimers and monomers. *J. Mol. Biol.* **271**, 266–277.

70. Hall, A., Ekiel, I., Mason, R.W., Kasprzykowski, F., Grubb, A. & Abrahamson, M. (1998) Structural basis for different inhibitory specificities of human cystatins C and D. *Biochemistry* **37**, 4071–4079.

71. Gudmundsson, G., Hallgrimsson, J., Jonasson, T.A. & Bjarnason, O. (1972) Hereditary cerebral hemorrhage with amyloidosis. *Brain* **95**, 387–404.

72. Löfberg, H., Grubb, A., Nilsson, E.K., Jensson, O., Gudmundsson, G., Blondal, H., Arnason, A. & Thorsteinsson, L. (1987) Immunohistochemical characterization of the amyloid deposits and quantification of pertinent cerebrospinal proteins in hereditary cerebral hemorrhage with amyloidosis. *Stroke* **18**, 431–440.

73. Abrahamson, M. & Grubb, A. (1994) Increased body temperature accelerates aggregation of the Leu→Gln mutant cystatin C, the amyloid-forming protein in hereditary cystatin C amyloid angiopathy. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1416–1420.

74. Olafsson, I. & Grubb, A. (2000) Hereditary cystatin C amyloid angiopathy. *J. Exp. Clin. Invest.* **7**, 70–79.

75. Gerhartz, B., Ekiel, I. & Abrahamson, M. (1998) Two stable unfolding intermediates of the disease-causing L68Q variant of human cystatin C. *Biochemistry* **37**, 17309–17317.

76. Ierovnik, E., Cimerman, N., Kos, J., Turk, V. & Lohner, K. (1997) Thermal denaturation of human cystatin C and two of its variants; comparison to chicken cystatin. *Biol. Chem. Hoppe-Seyler* **378**, 1199–1203.

77. Ekiel, I. & Abrahamson, M. (1996) Folding-related dimerization of human cystatin C. *J. Biol. Chem.* **271**, 1314–1321.

78. Jankowska, E., Banecki, B., Wiczk, W. & Grzonka, Z. (1999) Self-association of cystatin; in *Peptide Science – Present and Future* (Shomonishi, Y., ed.), pp. 654–656, Kluwer Academic Publishers, Dordrecht.

79. Oberg, K.A. & Fink, A.L. (1998) A new attenuated total reflectance Fourrier transform infrared spectroscopy method for the study of protein in solution. *Anal. Biochem.* **256**, 92–106.

80. Grzonka, Z., Kasprzykowski, F., Oldziej, S., Wiczk, W. & Grubb, A. (1993) Recent developments in the chemistry and biology of cystatins; in *Peptide Chemistry 1992* (Yanaihara, N., ed.) pp. 243–246, ESCOM, Leiden.

81. Dieckmann, T., Mitschang, L., Hofmann, M., Kos, J., Turk, V., Auerswald, E.A., Jaenicke, R. & Oschkinat, H. (1993) The structures of native phosphorylated chicken cystatin and of recombinant unphosphorylated variant in solution. *J. Mol. Biol.* **234**, 1048–1059.

82. Engh, R.A., Dieckmann, T., Bode, W., Auerswald, E.A., Turk, V., Huber, R. & Oschkinat, H. (1993) Conformational variability of chicken cystatin. Comparison of structures determined by X-ray diffraction and NMR spectroscopy. *J. Mol. Biol.* **234**, 1060–1069.

83. Alvarez-Fernadez, M., Abrahamson, M. & Su, X.-D. (2000) Crystal structure of cystatin D, an inhibitor of papain-like cysteine proteinases; in *Max-lab Activity Report 1999* (Andersen, J.N., Nyholm, S.L. & Sorensen, H., eds.) pp. 224–225.

84. Abrahamson, M., Mason, R.W., Hansson, H., Buttle, D.J., Grubb, A. & Ohlsson, K. (1991) Human cystatin C. Role of the N-terminal segment in the inhibition of human cysteine proteinases and in its inactivation by leucocyte elastase. *Biochem. J.* **273**, 621–626.

85. Grubb, A., Löfberg, H. & Barrett, A.J. (1984) The disuphide bridges of human cystatin C (γ-trace) and chicken cystatin. *FEBS Lett.* **170**, 370–374.

86. Barret, A.J. (1987) The cystatins: A new class of peptidase inhibitors. *Trends Biochem. Sci.* **12**, 193–196.

87. Kozak, M., Jankowska, E., Janowski, R., Grzonka, Z., Grubb, A., Alvarez-Fernadez, M., Abrahamson, M. & Jaskólski, M. (1999) Expression of a selonomethionyl derivative and preliminary crystallographic studies of human cystatin C. *Acta Crystallogr.* **D55**, 1939–1942.

88. Abrahamson, M., Dalbøge, H., Olafsson, I., Carlsen, S. & Grubb, A. (1988) Efficient production of native, biologically active human cystatin C by *Escherichia coli*. *FEBS Lett.* **236**, 14–18.

89. Matthews, B.W. (1968) Solvent content in protein crystals. *J. Mol. Biol.* **33**, 491–497.

90. Björck, L., Åkesson, P., Bohus, M., Trojnar, J., Abrahamson, M., Olafsson, I. & Grubb, A. (1989) Bacterial growth blocked by a synthetic peptide based on the structure of a human proteinase inhibitor. *Nature* **337**, 385–386.

91. Hall, A., Abrahamson, M., Grubb, A., Trojnar, J., Kania, P., Kasprzykowska, R. & Kasprzykowski, F. (1992) Cystatin C based peptidyl diazomethanes as cysteine proteinase inhibitors: Influence of the peptidyl chain length. *J. Enzyme. Inhib.* **6**, 113–123.

92. Kasprzykowski, F., Kasprzykowska, R., Plucińska, K., Kania, P., Grubb, A., Abrahamson, M. & Schalen, C. (2001) Peptidyl diazomethylketones as cysteine protease inhibitors structurally based upon the inhibitory centers of cystatins. *Pol. J. Chem.* (in press).

93. Leary, R., Larsen, D., Watanabe, H. & Shaw, E. (1977) Diazomethyl ketone substrate derivatives as active-site-directed inhibitors of thiol proteases. Papain. *Biochemistry* **16**, 5857–5861.

94. Tarnowska, M., Ołdziej, S., Liwo, A., Kania, P., Kasprzykowski, F. & Grzonka, Z. (1992) MNDO study of the mechanism of the inhibition of cysteine proteinases by diazomethyl ketones. *Eur. Biophys. J.* **21**, 217–222.

95. Johansson, L., Grubb, A., Abrahamson, M., Kasprzykowski, F., Kasprzykowska, R., Grzonka, Z. & Lerner, U.H. (2000) A peptidyl derivative structurally based on the inhibitory center of cystatin C inhibits bone resorption *in vitro*. *Bone* **26**, 451–459.

96. Barrett, A.J., Kembhavi, A., Brown, M.A., Kirschke, H., Knight, C.G., Tamai, M. & Hanada, K. (1982) L-*trans*-epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. *Biochem. J.* 201, 189–198.

97. Matsumoto, K., Mizoue, K., Kitamura, K., Tse, W.C., Huber, C.P. & Ishida, T. (1999) Structural basis of inhibition of cysteine proteases by E-64 and its derivatives. *Biopolymers* **51**, 99–107.

98. Kasprzykowski, F., Schalen, C., Kasprzykowska, R., Jastrzębska, B. & Grubb, A. (2000) Synthesis and antibacterial properties of peptidyl derivatives and cyclopeptides structurally based upon the inhibitory center of human cystatin C. *APMIS* **108**, 471–481.

99. Yamamoto, D., Matsumoto, K., Ohishi, H., Inoue, M., Kitamura, K. & Mizuno, H. (1991) Refined X-ray structure of papain E-64 complex at 2.1 Å resolution. *J. Biol. Chem.* **266**, 14771–14777.

100. Kozak, M., Kozian, E., Grzonka, Z. & Jaskólski, M. (1997) Crystalization and preliminary crystallographic studies of a covalent complex between papain an oxirane-based inhibitor; in *Peptides 1996* (Ramage, R. & Epton, E., eds.) pp. 551–552, European Peptide Society.

101. Drenth, J. & Jansonius, J.N. (1959) The unit cell of mercuripapain crystals. *Nature* **28**, 1718–1719.

102. Czaplewski, C., Grzonka, Z., Jaskólski, M., Kasprzykowski, F., Kozak, M., Politowska, E. & Ciarkowski, J. (1999) Binding modes of a new epoxysuccinyl-peptide inhibitor of cysteine proteases. Where and how do cysteine proteases ex-

press their selectivity? *Biochim. Biophys. Acta* **1431**, 290–305.

103. Drenth, J., Kalk, K.H. & Swen, H.M. (1976) Binding of chloromethyl ketone substrate analogues to crystalline papain. *Biochemistry* **15**, 3731–3738.

104. Pickersgill, R.W., Harris, G.W. & Garman, E. (1992) Structure of monoclinic papain at 1.6 Å resolution. *Acta Crystallogr.* **B48**, 59–67.

105. S*YBYL 6.1.* (1994) Tripos Inc., 1699 S. Manley Rd., St. Louis, MO 63144, U.S.A.

106. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. & Tasami, M. (1977) The protein data bank: A computer-based archive file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.

107. Case, D.A., Pearlman, D.A., Caldwell, J.W., Cheatham III, T.E., Ross, W.S., Darden, T., Merz, K.M., Stanton, R.V., Cheng, A., Vincent, J.J., Crowley, M., Ferguson, D.M., Radmer, R., Seibel, G.L., Singh, U.C., Wiener, P. & Kollman, P.A. (1997) *Amber 5.0,* University of California, San Francisco.

108. Bayly, C.I., Cieplak, P., Cornell, W.D. & Kollman, P.A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J. Phys. Chem.* **97**, 10269–10280.

109. Schroder, E., Phillips, C., Garman, E., Harlos, K. & Crawford, C. (1993) X-Ray crystallographic structure of a papain-leupeptin complex. *FEBS Lett.* **315**, 38–42.

110. Varughese, K.I., Ahmed, F.R., Carey, P.R., Hasnain, S., Huber, C.P. & Storer, A.C. (1989) Crystal structure of a papain-E-64 complex. *Biochemistry* **28**, 1330–1332.