Reprint from
Cysteine Proteinases and their Inhibitors
Editor: V. Turk

HUMAN STEFINS AND CYSTATINS : THEIR PROPERTIES AND STRUCTURAL RELATIONSHIPS

V.Turk, J.Brzin, B.Lenarčič, A.Šali

Department of Biochemistry, J. Stefan Institute, University E. Kardelj, Jamova 39, Ljubljana 61000, Yugoslavia

W. Machleidt

Institut fü Physiologische Chemie der Universität München, Goethestrasse 33, D-8000 München 2, FRG

Introduction

Although numerous protein inhibitors of proteinases have been described in many animals, plants and microorganisms, only recently has much attention been paid to cysteine proteinase inhibitors (CPIs). Their discovery offers a new insight into the processes in which they participate. At present, CPIs are believed to protect cells from innappropriate endogenous or external proteolysis and/or they could be involved in the control mechanisms responsible for intracellular or extracellular protein breakdown (1). Recently we have shown that some protein inhibitors of cysteine proteinases, such as chicken cystatin and human cystatin C and stefin B, are able to block virus replication when added to infected cells in culture (2,3).

CPIs were isolated from different tissues and body fluids of mammalian species including man (1, 4, 5). These inhibitors occur both intracellularly and extracellularly. According to their size they are described as low-M_r CPI or high-M_r CPIs. The amino acid sequence of known inhibitors show that they form a superfamily of CPIs (also named cystatins; see nomenclature and classification in this volume) which is constituted by at least three distinct families,the stefin family, the cystatin family and the kininogen family (6, Müller-Esterl et al., in this volume). Whereas members of the stefin and cystatin family are low-M_r proteins

with M_r of about 11,000-13,000 Da, kininogens (former λ-CPIs) are considerably larger molecules having M_r above 65,000 Da.

Low-M_r protein inhibitors, the stefins and the cystatins, are thermostable and are not glycoproteins. They are potent inhibitors of papain, cathepsin B, H and L, and dipeptidyl peptidase I (cathepsin C). They do not inhibit the cysteine proteinase bromelain. In addition, they do not inhibit serine, aspartic and metallo-proteinases. Mostly these inhibitors are isolated by a purification procedure which includes alkaline treatment and affinity chromatography on carboxymethylated papain.

In this paper we summarize the properties of human stefin A, human stefin B and human cystatin C. Their amino acid sequences are compared with sequences of other protein inhibitors of the stefin and cystatin family. A high degree of homology is observed throughout both families. Furthermore, the physical parametric approach to protein sequence comparison was used to find the structural resemblance of the stefins and cystatins, and their division into two families that have diverged during evolution.

Stefin family

Human stefin A, originally named human stefin, was isolated from polymorphonuclear granulocytes and characterized as an acidic protein with pI 4.65 (7). It is immunologically identical to human liver CPI-A (8) and to the acidic type of cysteine proteinase inhibitor (ACPI) isolated from human spleen (9). The inhibitor was found to be abundant in polymorphonuclear leucocytes in liver (10), thus confirming our successful isolation (7) which allowed us to purify this inhibitor in sufficient quantities to determine its primary structure (11, 12). The M_r of human stefin A is 11,006 Da, as calculated from the amino acid sequence (11) which is presented in Fig.1. The absence of Cys and Trp is evident. This was the first indication that the cysteine residue is not directly involved in inhibitory activity as was previously believed. The absence of cysteine and tryptophan was also observed in rat epidermal TPI (13), which is highly homologous to human stefin A as presented in Fig.1.

Human stefin A is a very potent inhibitor of papain , cathepsin H and L (17).

Human stefin B was isolated from human spleen in two forms with pI of about 5.9 and 6.5 (1, Lenarčič et al., in this volume). This inhibitor is similar or identical with an inhibitor that has been isolated from the same organ as a neutral type of cysteine proteinase inhibitor-NCPI (9) and from human liver (8), although there are some minor differences in their pI values. The inhibitor was isolated in an active monomer form of about 12,000 Da, as well as an inactive dimer form of about 24,000 Da, in almost equivalent amounts. The appearance of dimers in small quantities has already been reported(8,9). The inactive dimer can be converted to the inhibitory active monomer form under reducing conditions, indicating that the dimerization is the result of disulphide bond formation. These results are in agreement with similar behaviour of the monomer and dimer forms of rat liver TPI (14) and not with those for human liver cystatin B (15). Whether this observation has physiological importance has to be investigated in more detail.

The N-terminal sequence of human stefin B was determined (Lenarčič et al., in this volume, Lenarčič et al., in press) and is identical to the corresponding part of human liver cystatin B (15). Therefore we assume that both inhibitors are identical and they will be refered to as human stefin B throughout this paper. Stefin B contains only one cysteine residue at position 3, and no tryptophan (Fig.1). The carboxymethylated monomer form retains its inhibitory activity clearly showing that the thiol group is not essential for the activity. Human stefin B is highly homologous to rat liver TPI (16) with 79% identical amino acid residues. However, rat liver TPI contains two cysteine residues at position 3 and 64, Cys-64 being replaced by phenylalanine in human stefin B. This is additional evidence that cysteine residues are not important for the inhibitory activity.

Human stefin B is a potent inhibitor of papain, cathepsin H and L and a weaker inhibitor of cathepsin B (17).

Cystatin family

Human cystatin C was isolated from serum of patients with autoimmune diseases as an alkaline low-M_r protein inhibitor (18, 19). The inhibitor occurs in several multiple forms, ranging from pI 7.8-9.1, possibly as the result of the degradation of the amino terminal part of the molecule, producing forms lacking one or more amino acid residues. When its partial amino acid sequence was determined (18,19), we discovered that this inhibitor is identical to human γ-trace or post γ-globulin, a basic microprotein, of hitherto unknown biological function, isolated from the urine of patients with renal failure (20). Due to its high degree of homology with chicken cystatin, which we sequenced almost at the same time (18), we proposed that the newly discovered inhibitor of cysteine proteinases should be named human cystatin (18,19), and soon after the name human cystatin C was suggested (17). It was reported previously that γ-trace is a componentof amiloid fibrils in patients suffering from hereditary cerebrovascular amyloidosis, thus indicating its very important biological role in metabolic processes (21). Very recently a variant of γ-trace was found, which has an amino acid substitution (Gln for Leu) at position 68 (cystatin C numbering), which is near the proposed active site of the CPIs (22). It is suggested that this mutation might lead to the production of this abnormaly degraded bound and/or precipitated protein.

Human cystatin S was isolated from human saliva as an acidic inhibitor with pI 4.68 and M_r of 13,255 Da, as determined from the amino acid sequence (23,24). Recent studies show that human saliva contains several molecular variants of cystatin S (forms S5 and S7) as a result of the heterogeneity in the N-terminal part of the inhibitor (Isemura et al., in this volume).

Very recently a new inhibitor, cystatin SN with a pI of 7.5 was isolated also from human saliva and sequenced (25). All investigated inhibitors from human saliva are highly homologous to cystatin C (19,20) and chicken cystatin(18,26).

Amino acid sequences of the cystatin family are presented in Fig.2.

Sequence homologies among the stefin abd cystatin family

The amino acid sequences of known low-M_r CPIs clearly show that according to the criterion of sequence homology they fall into two families: the stefin and the cystatin family. Both families (sequences shown in Fig.1, Fig.2 and Fig.3) together with the kininogen family belong to a single superfamily (6).

The members of the stefin family (Fig.1) are proteins composed of 98-103 amino acid residues with M_r of about 11,000 Da. All inhibitors of human and rat origin have a N-terminal methionine, which is N-acylated in three of the four. Probably the methionine residues represent the points of the initiation of translation and therefore this family of inhibitors does not have an N-terminal signal sequence for translocation into extracellular space. In contrast, members of the cystatin family do not contain methionine as the N-terminal amino acid residue.

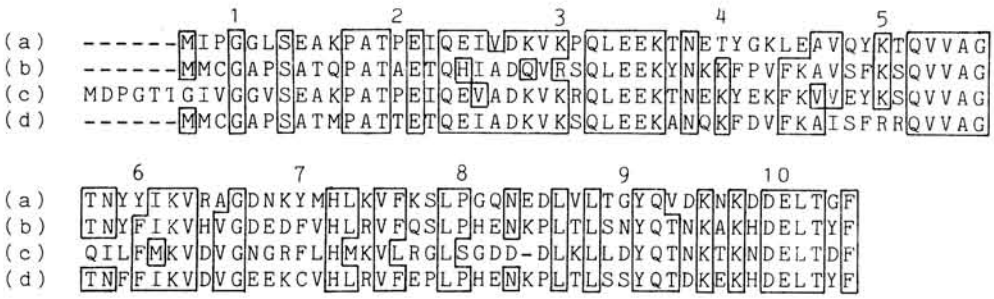

Fig.1. Comparison of the sequences of members of the stefin family : a) human stefin A, b) human stefin B, c) rat epidermal TPI, d) rat liver TPI. Residues identical in at least three of the four sequences are boxed.

Members of the stefin family share the highly conserved pentapeptide Gln-Val--Val-Ala-Gly region (residues 52-56, rat epidermal TPI numbering in Fig.1), presumably part of the proposed reactive site (1). Additional homologous amino acid residues are present in the other parts of the molecule. The members of the cystatin family (Fig.2) also show several regions with a high degree of homology (residues 49-75, residues 102-106 and dipeptide Cys-Gln at the C-termini; human cystatin C numbering in Fig.2). The most striking feature is the conservation of all four cysteine residues in all members of the cystatin family, and two cystatin segments of the kininogen molecule. It was reported (27) that the four

```
                1           2           3           4           5
(a)  SSPGKPPRLVGGPMDASVEEEGVRRALDFAVGEYNKASNDMYHSRALQVVRARKQIVAG
(b)  ----ENRIIPGGIYDADLNDEWVQRALHFAISEYNKATEDEYYRRPLQVLRAREQTFGG
(c)  --------IIPGGIYNADLNDEWVQRALHFAISEYNKATKDDYYRRPLRVLRARQQTVGG
(d)  ---EEDRIIEGGIYDADLNDERVQRALHFVISEYNKATEDEYYRRLLRVLRAREQIVGG
(e)  --------RLLGGLMEADVNEEGVQEALSFAVSEFNKRSNDAYQSRVVRVVRARKQVVSG
(f)  --SEDRSRLLGAPVPVDENDEGLQRALQFAMAEYNRASNDKYSSRVVRVISAKRQLVSG

          6           7           8           9          10          11          12
(a)  VNYFLDVELGRTTCTKTQPNLDNCPFHDQPHLKRKAFCSFQIYAVPWQGTMTLSKSTCQDA
(b)  VNYFFDVEVGRTICTKSQPNLDTCAFHEQPELQKKQLCSFEIYEVPWEDRMSLVDSRCQEA
(c)  VNYFFDVEVGRTICTKSQPNLDTCAFHEQPELQKKQLCSFEIYEVPWENRRSLVKSRCQES
(d)  VNYFFDIEVGRTICTKSQPNLDTCAFHEQPELQKKQLCSFQIYEVPWEDRMSLVDSRCQEA
(e)  MNYFLDVELGRTTCTKSQANLDSCPFHNQPHLKREKLCSFQVYVVPWMNTINLVKFSCQD-
(f)  IKYILQVEIGRTTCPKSSGDLQSCEFHDEPEMAKYTTCTFVVYSIPWLNQIKLLESKCQ--
```

Fig.2. Comparison of the sequences of the members of the cystatin family: a) a) human cystatin C, b) human cystatin S7, c) human cystatin SN, d) human cystatin S5, e) bovine colostrum cystatin, f) chicken cystatin. All common residues within the cystatin family are boxed.

cysteines in human cystatin C and chicken cystatin form two disulphide bonds (Cys71-Cys81 and Cys95-Cys115), and at least one disulphide bond is essential for inhibitory activity. The same disulphide structure was confirmed for bovine colostrum CPI, which belongs to the cystatin family(28).

The alignment of sequences of the stefin and cystatin family in Fig.3 shows that only eight amino acid residues are common to both families. However,only two residues Gln55 and Gly59 (rat epidermal TPI numbering in Fig.3), are common to all the sequences of the stefin family, the cystatin family and two inhibitory segments of the kininogen molecule (6). These residues might have functional importance because they are absent from the first, non-inhibitory cystatin segment of kininogen (29).

In order to show structural similarities between individual members of the stefin and the cystatin family, we have constructed the hydropathy profiles according to Kyte and Doolittle (30), using a moving segment of 6 residues. Hydropathy profiles of the stefin family (Fig.4) and the cystatin family (Fig.5) indicate, that the members of each family should have very similar tertiary structures. The most pronounced features are hydrophilicity of the amino and carboxy terminal regions and hydrophobicity of the central region of the molecule. The most conserved region (residues 49-69), assumed to interact with the target enzymes,

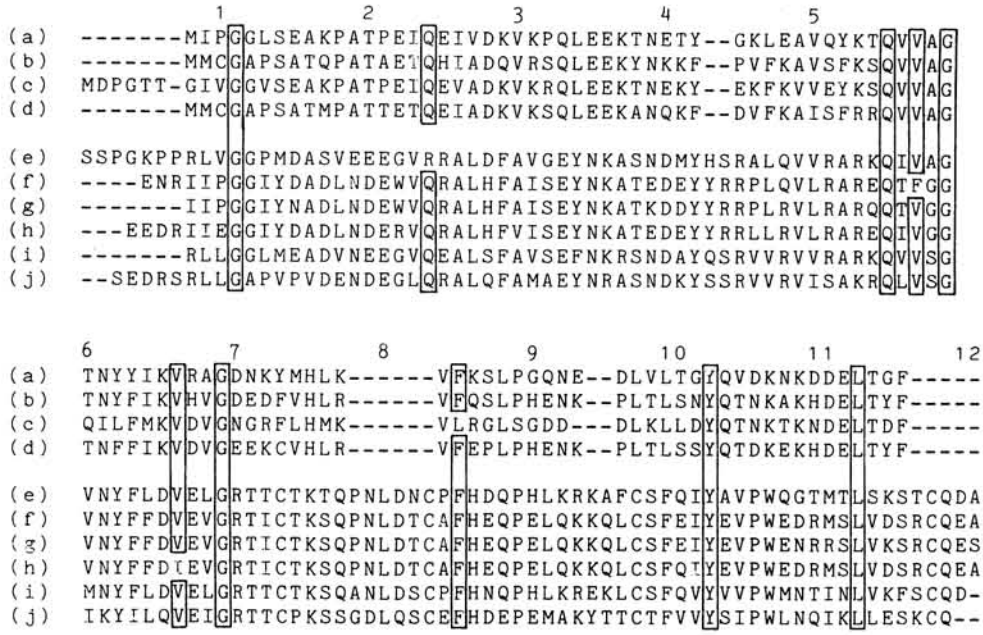

Fig.3 Comparison of the sequences of the members of the stefin and cystatin family : a)human stefin A, b) human stefin B, c) rat epidermal TPI, d) rat liver TPI, e) human cystatin C, f) human cystatin S7, g) human cystatin SN, h) human cystatin S5, i) bovine colostrum cystatin, j) chicken cystatin. Common residues (nine of ten) are boxed.

is hydrophobic. A more detailed knowledge of the tertiary structure is expected to be obtained from X-ray structure analysis. The first inhibitor of CPIs, chicken cystatin, has already been cristallized (31), and would probably help to establish the mode of comlpex formation with cysteine proteinases.

Comparison of the known sequences of different CPIs from the stefin, cystatin and kininogen family enabled the construction of a scheme for the evolution of mammalian CPIs (6). This model suggests that the diversity of these proteins has evolved from two ancestral units forming the stefins and cystatins. In addition, the physical parametric approach to protein sequence comparison (32) has been used to make a structural comparison based on amino acid sequences between the CPIs of the stefin and cystatin family whose sequences are known. The

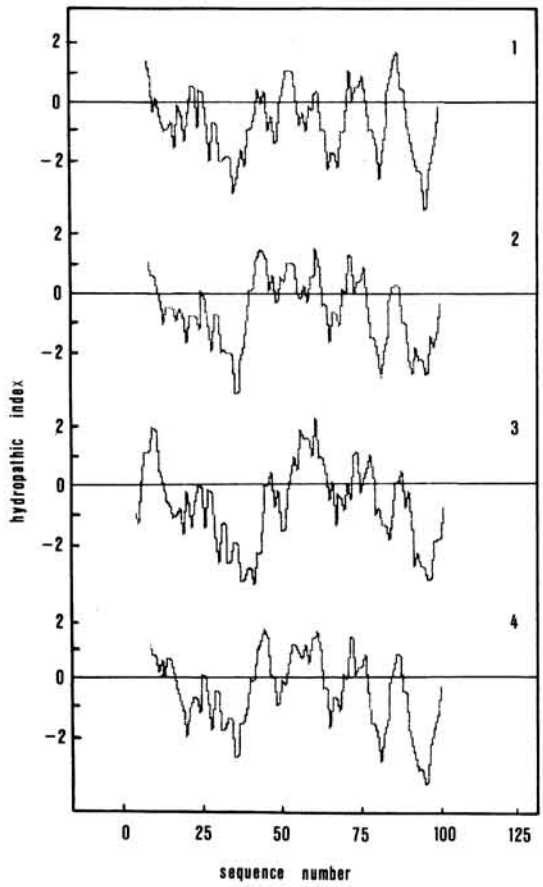

Fig.4. Hydropathy profiles of members of the stefin family : 1) human stefin A, 2) human stefin B, 3) rat epidermal TPI, 4) rat liver TPI.

comparison is based on the physical characteristics thought to determine the three dimensional folding of a given polypeptide sequence, and may be applicable for distantly related proteins which maintain similar structural folds without any apparent sequence homology. Such a dendrogram was constructed from the difference matrix and shows the development by divergent evolution of the

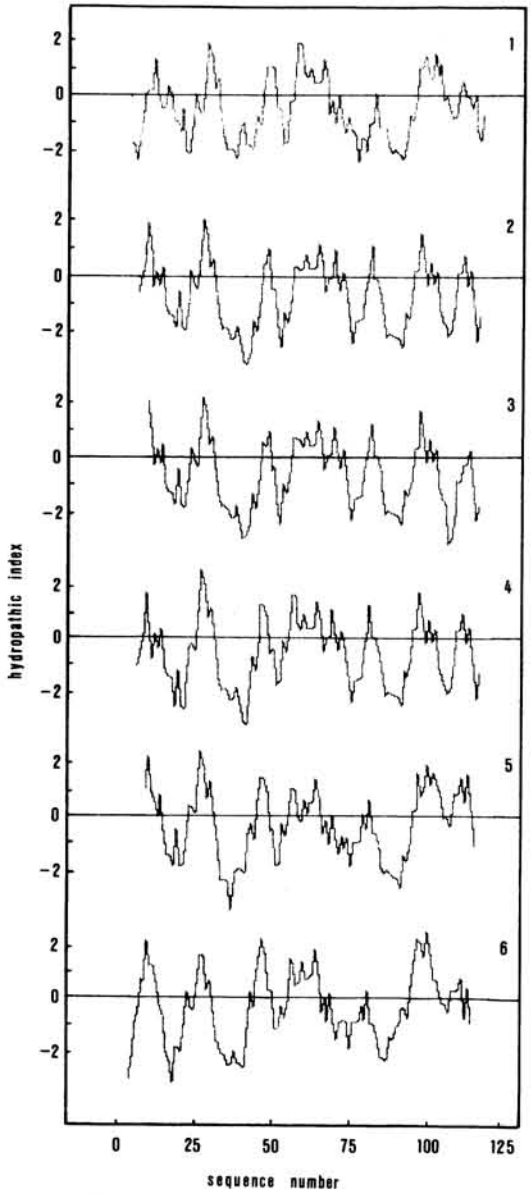

Fig.5. Hydropathy profiles of members of the cystatin family : 1) human cystatin C, 2) human cystatin S7, 3) human cystatin SN, 4) human cystatin S5, 5) bovine colostrum cystatin, 6) chicken cystatin.

stefin and cystatin family. It is evident that the most closely related inhibitors so far sequenced are human stefin B and rat liver TPI (Fig.6, see also Lenarčič et al. in this volume).

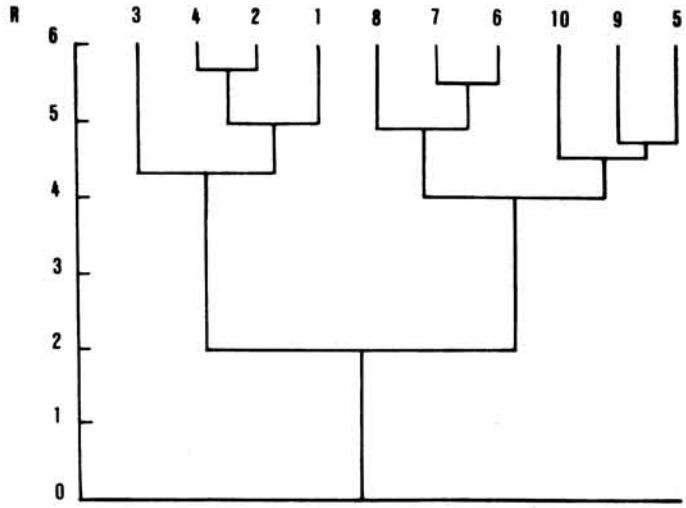

Fig.6. Dendrogram constructed from the difference matrix. Cophenetic correlation coefficient: 0.9634. R = maximal summed cross correlation coefficient. 1) human stefin A, 2) human stefin B, 3)rat epidermal TPI, 4)rat liver TPI, 5)human cystatin C, 6)human cystatin S7, 7)human cystatin S5, 8)human cystatin SN, 9)bovine colostrum cystatin, 10)chicken cystatin.

CONCLUSION

Progress made within tha last two years towards elucidating the primary structure of cysteine proteinase inhibitors, together with the discovery that kininogens are also inhibitors of cysteine proteinases, made it possible to classify these groups of proteins into three families of CPIs: the stefins, the cystatins and the kininogens. All three families belong to the same superfamily of cystatins. Sequence homology of CPIs enables us to construct a dendrogram which shows the evolutionary relatedness of these proteins. Further knowledge of their structure and the mechanism of action is necessary to understand their biological role in the normal and pathological processes in which they participate.

Discovery of new inhibitors also involves questions of nomenclature. Once amino acid sequence is established, a systematic family name should be maintained, preceeded by a prefix indicating origin, and followed by a letter or number for subdivision within the same species.

For better understanding the list of individual inhibitors so far sequenced is presented and grooped into logical family groups.

THE SUPERFAMILY OF CYSTATINS

Stefin family (family 1)

human stefin A (cystatin A)
human stefin B (human CPI B or cystatin B)
rat stefin A (rat epidermal TPI or rat cystatin α)
rat stefin B (rat liver TPI or rat cystatin β)

Cystatin family (family 2)

human cystatin C (β- trace or post β - globulin)
human cystatin S
human cystatin SN
human cystatin S5
human cystatin S7
chicken cystatin
bovine cystatin

Kininogen family (family 3)

human high-M_r kininogen
bovine high-M_r kininogen
rat high-M_r kininogen
human low-M_r kininogen
bovine low-M_r kininogen
rat low-M_r kininogen
rat T-kininogen

Acknowledgement

This work was supported by grants from the Research Council of Slovenia (C1--0114-106), Yugoslavia, from the Sonderforschungsbereich 0207 of the University of Munich, from the International Bureau of the Kernforschungsanlage Jülich GmbH, and from the Bundesministerium für Forschung und Technologie, FRG (grant PTB 8651). The investigations were supported also by the NSF Grant No. F7F030Y.

References

1.Turk,V., J.Brzin, B.Lenarčič, P.Ločnikar, T.Popovič, A.Ritonja, J.Babnik, W.Bode, W.Machleidt. 1985. In: Intracellular Protein Catabolism (E.A. Khairallah, J.S.Bond and J.W.C.Bird, eds.). A.R.Liss, New York, pp.91-103.

2.Korant,B., J.Brzin, V.Turk. 1985. Biochem.Biophys.Res.Commun. 127, 1072-1076.

3.Korant,B., T.Towatary, L.Ivanoff, J.Brzin, B.Lenarčič, V.Turk. 1986. UCLA Symposia on Molecular and Cellular Biology (in press).

4.Barrett,A.J. 1985. In: Intracellular Protein Catabolism (E.A.Khairallah, J.S.Bond and J.W.C.Bird, eds.). A.R.Liss, New York, pp. 105-116.

5.Turk,V., J.Brzin, M.Kopitar, M.Kotnik, B.Lenarčič, T.Popovič, A.Ritonja, M.Trstenjak, B.Rozman, W.Machleidt. 1986. In : Proteinases in Inflammation and Tumor Invasion (H.Tschesche, ed.). Walter de Gruyter, Berlin (in press)

6.Müller-Esterl,W., H.Fritz, J.Kellermann, F.Lottspeich, W.Machleidt, V. Turk. 1985. FEBS Lett. 191, 221-226.

7.Brzin,J., Kopitar,M., V.Turk. 1983. Hoppe-Seyler´s Z.Physiol.Chem. 364, 1475-1480.

8.Green , G.D.J.,A.A.Kembhavi, M.E.Davies, A.J.Barrett. 1984. Biochem.J. 218, 939-946.

9.Järvinen,M., A.Rinne. 1982. Biochim.Biophys.Acta 708, 210-217.

10.Davies,M.E., A.J.Barrett. 1984. Histochemistry 80, 373-377.

11.Machleidt,W., Borchart,U., H.Fritz, J.Brzin, A.Ritonja, V.Turk. 1983. Hoppe-Sayler´s Z.Physiol.Chem. 364, 1481-1486.

12.Turk,V., J.Brzin, M.Kopitar, I.Kregar, P.Ločnikar, M.Longer, T.Popovič, A.Ritonja, Lj.Vitale, W.Machleidt, T.Giraldi, G.Sava. 1983. In: Proteinase Inhibitors : Medical and Biological Aspects (N.Katunuma, H.Umezawa, H. Holzer, eds.). Jap.Sc.Soc.Press, Springer Verlag, Tokyo-Berlin,pp.125-134.

13.Takio,K., E.Kominami, Y.Bando, N.Katunuma, K.Titani. 1984. Biochem. Biophys.Res.Commun. 121, 149-154.

14.Wakamatsu,N., E.Kominami, K.Takio, N.Katunuma. 1984. J.Biol.Chem. 259, 13832-13838.

15.Ritonja,A., W.Machleidt, A.J.Barrett. 1985.Biochem.Biophys.Res.Commun. 131, 1187-1192.

16.Takio,K., E.Kominami, N.Wakamatsu, N.Katunuma, K.Titani. 1983. Biochem.Biophys.Res.Commun. 115, 902-908.

17.Barrett, A.J., M.E.Davies, A.Grubb. 1984. Biochem.Biophys.Res.Commun. 120, 631-636.

18.Turk,V., J.Brzin, M.Longer, A.Ritonja, M.Eropkin, U.Borchart, W.Machleidt. 1983.Hoppe-Seyler's Z.Physiol.Chem.364, 1487-1496.

19.Brzin,J., T.Popović, V.Turk, U.Borchart, W.Machleidt. 1984. Biochem. Biophys.Res.Commun. 118, 103-109.

20.Grubb,A., H.Löfberg. 1982. Proc.Natl.Acad.Sci.USA 79, 3024-3027.

21.Cohen,D.H.,H.Feiner, O.Jensson, B.Frangione. 1983.J.Exp.Med. 158, 623-628.

22.Ghiso,J., O.Jensson, B.Frangione. 1986. Proc.Natl.Acad.Sci.USA (in press)

23.Isemura,S., E.Saitoh, S.Ito, M.Isemura, K.Sanada. 1984. J.Biochem. 96, 1311-1314.

24.Isemura,S., E.Saitoh, K.Sanada.1984. J.Biochem. 96, 489-498.

25.Isemura,S., E.Saitoh, K.Sanada. 1986. FEBS Lett. 198,145-149.

26.Schwabe,C., A.Anastasi, H.Crow, J.K.McDonald, A.J.Barrett. 1984. Biochem.J. 217, 813-817.

27.Grubb,A., H.Löfberg, A.J.Barrett. 1984. FEBS Lett. 170, 370-374.

28.Hirado,M., S.Tsunasawa, F.Sakiyama, M.Niinobe, S.Fujii. 1985. FEBS Lett. 186, 41-45.

29.Salvesen,G., C.Parkes, M.Abrahamson, A.Grubb, A.J.Barrett. 1986. Biochem.J. 234, 429-434.

30.Kyte,J., Doolittle.1982. J.Mol.Biol. 157, 105-132.

31.Bode,W., J.Brzin, V.Turk. 1985. J.Mol.Biol. 181, 331-332.

32.Argos,P., M.Hanei, J.M.Wilson, W.N.Kelley. 1983. J.Biol.Chem. 258, 6450-6457.