Volume 243, number 2, 186–192 FEB 06717 January 1989

Synthesis of a (desSer1 Ile29 Leu89) chicken cystatin gene, expression in *E. coli* as fusion protein and its isolation

Ennes A. Auerswald, Gabriele Genenger, Irmgard Assfalg-Machleidt*, Janko Kos+ and Wolfram Bode°

*Abteilung für Klinische Chemie und Klinische Biochemie in der Chirurgischen Klinik Innenstadt, Ludwigs-Maximillians-Universität München, Nußbaumstr. 20, D-8000 München 2, *Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München, Goethestr. 33, D-8000 München 2, FRG, +Odsek za biokemijo, Institut Jozef Stefan, Univerza E. Kardelja, Jamova 39, YU-61000 Ljubljana, Yugoslavia* and *°Max-Planck-Institut für Biochemie, D-8033 Martinsried, FRG*

Received 21 November 1988; revised version received 1 December 1988

A synthetic gene coding for the cysteine proteinase inhibitor (desSer1 Ile29 Leu89) chicken cystatin was cloned and expressed in *E. coli*. The gene was assembled from 12 oligonucleotides and inserted into vector pUC 8. Expression as fusion protein was performed in a temperature-inducible *E. coli* system. The expression product was synthesized as 20% of total *E. coli* protein. The fusion protein was purified, the chicken cystatin homologue was split off with CNBr and the N-terminal sequence confirmed up to position 37. The properties of the purified material correspond to those of natural chicken cystatin. The recombinant cystatin variant binds anti-chicken cystatin IgG, is inhibitorily active and displays K_i values with papain and with cathepsin B similar to those determined for natural chicken cystatin.

Cysteine proteinase; Enzyme inhibitor; Cystatin; Synthetic gene; Protein engineering; (Chicken)

1. INTRODUCTION

Chicken cystatin of M_r 13000 is one of the best investigated tight binding inhibitors of cysteine proteinases [1]. It was originally isolated from egg white [2]; later on two forms of different p*I* and truncated forms were characterized [3–5]. The dissociation constants of their papain complexes were measured ([6] Machleidt, W. et al., submitted). The amino acid sequence of one form was determined [3,7], and Gly9 Ala10 were suggested to represent the putative reactive site residues [8]. Very recently the X-ray crystal structure of chicken cystatin was solved and the possible mode of interaction with papain elucidated [9]. The detailed knowledge of its structure and binding mode as well as the expected therapeutic potential of cysteine proteinase inhibitors made a recombinant production of this interesting molecule very attractive. Therefore a synthetic master gene was designed, synthesized, cloned, expressed and the purified product biochemically characterized.

Correspondence address: E.A. Auerswald, Abteilung für Klinische Chemie und Klinische Biochemie in der Chirurgischen Klinik Innenstadt, Universität München, Nußbaumstr. 20, D-8000 München 2, FRG

Abbreviations: aa, amino acids; MS-2 pol, RNA bacteriophage MS-2 polymerase; PAGE, polyacrylamide gel electrophoresis

2. MATERIALS AND METHODS

2.1. *Oligonucleotide synthesis and fragment assembly*

The 12 oligonucleotides were synthesized by the phosphite triester method [10] using the phoshoramidite chemistry [11] on a Pharmacia gene assembler. The material was purified by denaturing polyacrylamide gel electrophoresis. The hybridisation of oligonucleotides (100 pmol each) and their ligation into segments A and B is outlined in fig.1 and was performed as described elsewhere [12,13].

2.2. *Construction of cloning and expression vectors*

Segments A and B were purified on 2% agarose gels after cleavage with *Pst*I and ligated into pUC 8 vectors cut with

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

*Eco*RI/*Pst*I and *Pst*I/*Hin*dIII, respectively (molar ratio of vector:fragment, 1:5). Transformed *E. coli* RR1 delta M15 cells [14] were selected after plasmid preparation [15], restriction analysis and DNA sequence analysis [16]. The complete gene was constructed by isolating segment A from pGG 1.6.1 and ligating it into the *Eco*RI/*Pst*I cleaved vector pGG 5.1.2 harbouring segment B. Out of several transformants a construct with the correct DNA sequence of the (desSer1 Ile29 Leu89) chicken cystatin gene was selected (pGG 8.1.1). The gene incorporated in an *Eco*RI/*Bam*HI fragment was purified and ligated into expression vector pEx 31b [17]. As expression host we used *E. coli* CGSC 6769 (Central Genetic Stock Center, Yale) which was transformed with pcI 857 to provide the system with the heat labile repressor cI 857, which is necessary for heat induction [18]. Analysis of the correct insertion was done by direct screening for fusion proteins after induction of the *E. coli* cells [13].

2.3. *Expression and isolation of the chicken cystatin homologue*

The expression of the fusion protein, its isolation and the CNBr cleavage of the fusion protein were performed as described elsewhere [13,19], with modifications for heat induction (1 h, 42°C and 3 h, 32°C). The reaction mixture containing the uncleaved fusion protein, the chicken cystatin homologue, the fusion protein part of MS-2 pol and the acidolytic fragment of the MS-2 pol part (resulting from the cleavage of the acid sensitive Asp-Pro bond) were fractionated by gel filtration on Sephadex G-50 and these fractions were analysed separately.

2.4. *Biochemical and immunological characterization*

Natural chicken cystatin was prepared as described [20]. This material was used to prepare antibodies in rabbits [21]. From antisera, IgGs were immunoselected by affinity chromatography [22]. For SDS-PAGE [23], Western blot analysis [24,25] and amino acid sequencing [26] recombinant and natural material were used. For detection of inhibitory activity purified fusion protein and (desSer1 Ile29 Leu89) chicken cystatin were refolded during stepwise dilution from 6 M to 2 M urea followed by passage through a small Sephadex G-25 column (NAP-10, Pharmacia) equilibrated with the assay buffer [13]. Fluorometric assays were used for titrations with papain and cathepsin B. The K_i estimations with papain (EC 3.4.22.2, type III, Sigma, repurified by covalent chromatography on mercurial agarose [27]) and with human cathepsin B (EC 3.4.22.1; kindly provided by T. Popovic) were done as described elsewhere [13,28,29]. Natural highly purified chicken cystatin used as reference was only the long form of the inhibitor ([5] Machleidt, W. et al., submitted).

3. RESULTS

3.1. *Design, gene assembly and cloning*

A synthetic gene coding for a homologue of chicken cystatin was designed on the basis of its amino acid sequence [3,7] and its utilization as recombinant gene product for protein engineering. The structure of the gene and the strategy of synthesis are outlined in fig.1. The oligonucleotide building blocks with the codon usage for highly expressed proteins in *E. coli* were chosen [30]. Unique cleavage sites were selected by help of the UWGCG sequence analysis software [31]. The strategy for cloning and expression of the gene included the incorporation of an *Eco*RI site, an ATG start codon at the 5'-end, two stop codons as well as a *Bam*HI and a *Hin*dIII site at the 3'-end. The internal Met residues at positions 29 and 89 were substituted by Ile and Leu, respectively, to allow the release of an intact inhibitor by CNBr cleavage after the ATG coded Met residue linking the MS-2 pol part with the inhibitor. The following residue, Ser1, the first amino acid of the natural inhibitor was deleted because it is known that Met-Ser peptide bonds are cleaved poorly by CNBr [32]. The substitutions should alter neither structure nor function because both methionines are replaced in some of the related cystatins by Ile29 and by Leu89 [1]. The first amino acid of chicken cystatin, Ser1, is assumed not to be structurally or functionally important [9]. Based on experience obtained in construction of genes for aprotinin and stefin B [12,13], the gene was assembled in a modular way as shown in figs 1 and 2. Both purified segments A and B were cloned separately into vector pUC 8; among several transformants the vectors pGG 1.6.1 and pGG 5.1.2 were selected after DNA sequencing. The correct fragments were used for construction of pGG 8.1.1 harbouring the complete gene. Additional transformants containing certain deletions were found and will be described elsewhere.

3.2. *Expression of (desSer1 Ile29 Leu89) chicken cystatin*

For expression of the chicken cystatin homologue gene a temperature inducible *E. coli* expression system with the vectors pEx 31b, pcI857 and *E. coli* strain CGSC 6769 was chosen [13,17,19]. After induction and analysis of total *E. coli* proteins by SDS-PAGE the resulting *E. coli* strain 1083 was selected, expressing the fusion of MS-2 pol–chicken cystatin homologue. The main protein band migrating at an M_r of 24000 corresponds to 25% of total *E. coli* proteins (fig.3). The fusion protein aggregated as inclusion bodies and was immunologically detected after Western blotting with antibodies against native chicken cystatin (fig.3).

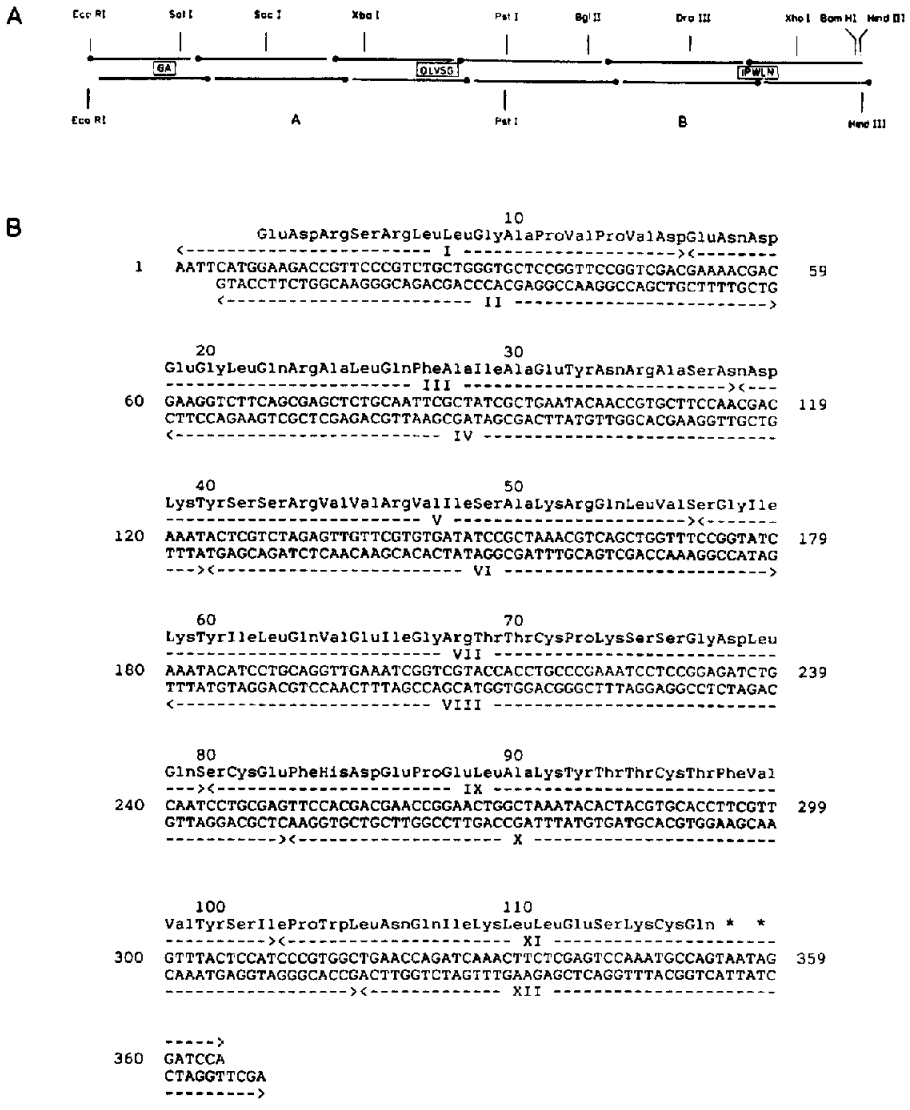

Fig.1. (A) Design of the synthetic chicken cystatin homologue gene. The boxed amino acids are believed to be part of structurally important sites. (B) DNA and amino acid sequences of (desSer1 Ile29 Leu89) chicken cystatin.

3.3. *Isolation and characterization*

From 3.5 l fermentation broth (10 g *E. coli* wet wt) about 1.5 g inclusion bodies (wet wt) were isolated and purified by gel filtration under reducing (10 mM β-mercaptoethanol) and denaturing conditions (6 M urea) on a Sephacryl S-300 column (about 200 mg). After CNBr cleavage the fragments: MS-2 pol fusion protein of 216 aa, chicken cystatin homologue of 115 aa, MS-2 pol of 101 aa and the acidolytic MS-2 pol fragment of 66 aa (resulting from the cleavage of the acid labile Asp35–Pro36 bond) were detected (fig.3) and separated on a Sephadex G-50 column. Fractions containing essentially pure (desSer1 Ile29 Leu89) chicken cystatin, contaminated with minor amounts of fusion protein, were isolated (fig.3). The material was used for amino acid sequencing, Western blotting (fig.3) and inhibition studies (table 1). The N-terminal amino acid sequence was confirmed up to position 37 including Ile29

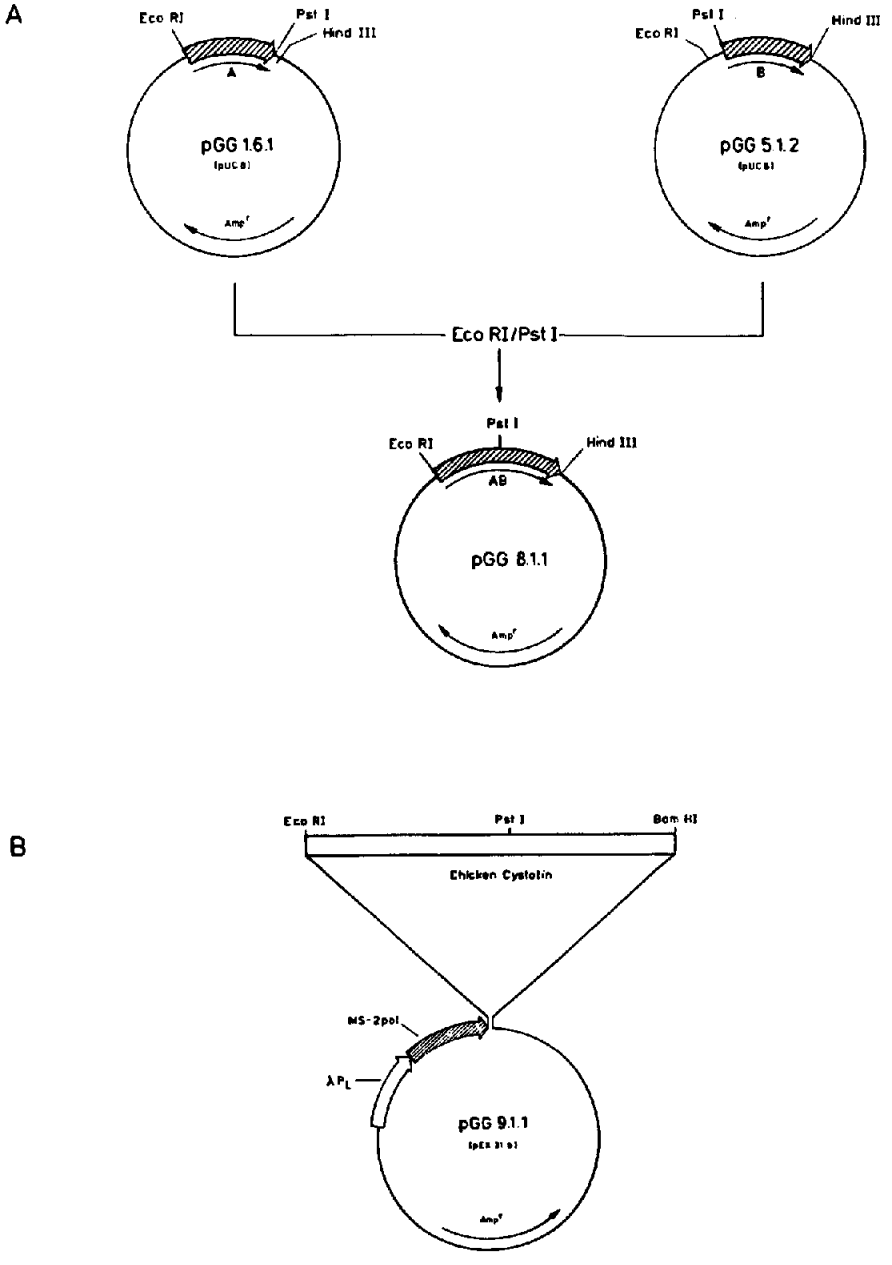

Fig.2. (A) Construction of cloning vectors. (B) The expression vector pGG 9.1.1. Only relevant restriction sites are shown.

substituted for Met (table 1). Contamination with MS-2 pol was less than 3%. The substituted amino acid residue Leu89 was identified in the expected position by sequencing unseparated tryptic fragments that had been selectively immobilized via their lysyl side chains ([26], not shown). After partial refolding by passage through a Sephadex G-25 column 37% of the fusion protein and 40% of (desSer1 Ile29 Leu89) chicken cystatin quantitated by A_{280} determination were found to be inhibitorily active by titration with papain (table 2). The estimated K_i values for the inhibition of pa-

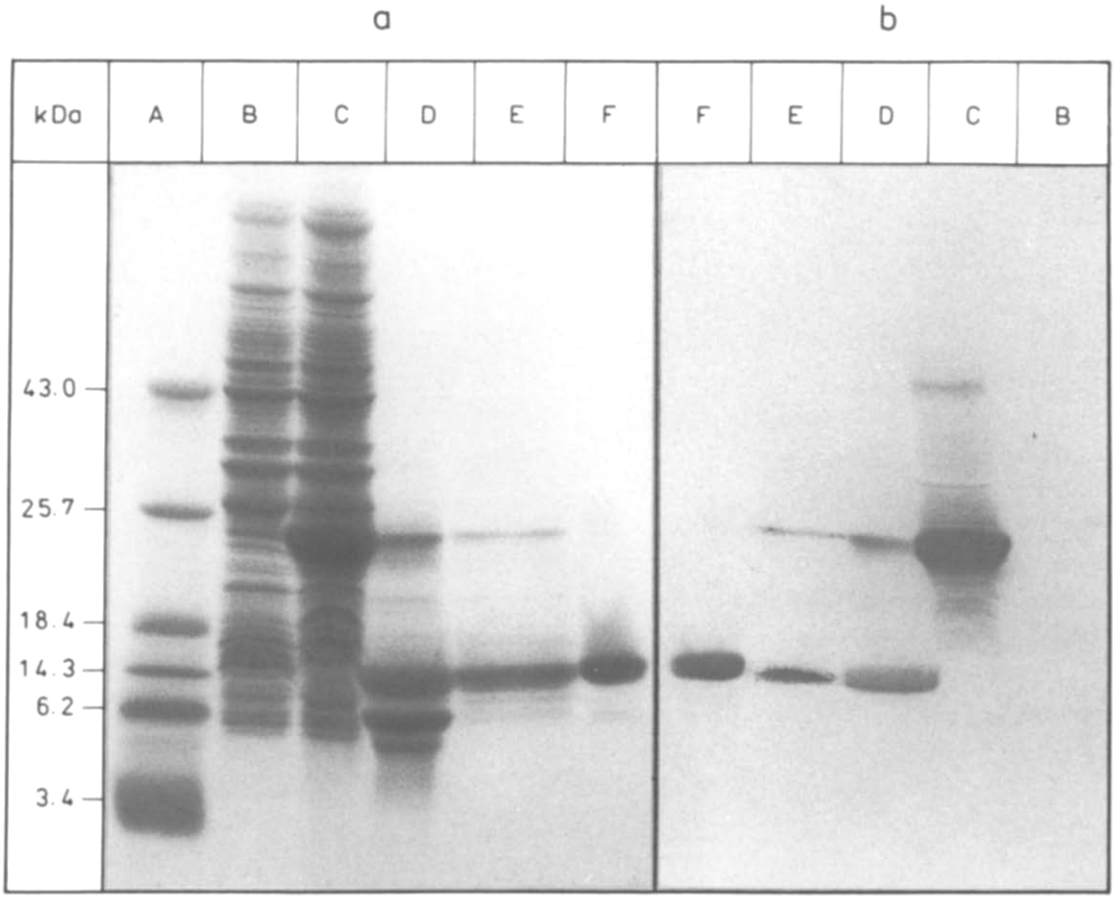

Fig.3. Electrophoretic analysis by (a) SDS-PAGE (5–20%) and by (b) Western blot with immunoprint. (A) Low molecular mass standards; (B) *E. coli* lysate with pEx 31b (4×10^8 cells); (C) *E. coli* lysate with pGG 9.1.1 (4×10^8 cells); (D) cleavage products after CNBr treatment (about 1 mg wet wt); (E) (desSer1 Ile29 Leu89) chicken cystatin after gel filtration (20 μl sample, $A_{280} = 0.9$) and (F) natural chicken cystatin (20 μl sample, $A_{280} = 0.9$).

pain and human cathepsin B are very similar to those obtained with natural chicken cystatin. These assays were done in parallel for natural and recombinant material: the values closely resemble those reported previously [1]. No significant differences were found between the inhibition constants of (desSer1 Ile29 Leu89) chicken cystatin and of the fusion protein (table 2).

4. DISCUSSION

We have cloned a synthetic gene of chicken cystatin and expressed it in *E. coli*, as a fusion protein together with 101 amino acids of MS-2 pol. We have further shown that modified (desSer1 Ile29 Leu89) chicken cystatin can be recovered from the recombinant fusion protein after cleavage with CNBr in a biologically active form. Within the expected limits of error, the K_i values for the inhibition of papain and human cathepsin B are identical with those of the natural inhibitor. With the isolation procedure described here about 5 mg of pure inhibitor was obtained per litre *E. coli* broth. Scaling-up and improvement of isolation should further increase the yield. The selected approach enables us to produce variants which according to the X-ray crystal structure and the suggested binding mode [9] might be critical either for the structure of this molecule or for effective binding to cysteine proteinases. These variants will further be used to study the refolding pathway and the significance of various structural components

Table 1

N-terminal amino acid sequence of (desSer1 Ile29 Leu89) chicken cystatin

Step	Residue	Yield[a] (nmol)	Step	Residue	Yield[a] (nmol)
1	X	[b]	20	Leu	0.90
2	Asp	3.02	21	Gln	0.45
3	Arg	1.90	22	Arg	0.59
4	Ser	0.54	23	Ala	0.53
5	Arg	0.84	24	Leu	0.67
6	Leu	1.15	25	Gln	0.31
7	Leu	1.33	26	Phe	0.39
8	Gly	0.65	27	Ala	0.41
9	Ala	0.97	28	Ile	0.39
10	Pro	0.54	29	Ala	0.45
11	Val	1.19	30	Glu	0.17
12	Pro	0.39	31	Tyr	0.32
13	Val	0.95	32	Asn	0.24
14	Asp	0.68	33	Arg	0.40
15	Glu	0.51	34	Ala	0.30
16	Asn	0.59	35	Ser	0.16
17	Asp	0.77	36	Asn	0.27
18	Glu	0.50	37	Asp	0.28
19	Gly	0.29			

[a] Raw yields of amino acid phenylthiohydantoin derivatives as determined by on-line HPLC without correction for background and overlap

[b] Due to the applied method [26], the N-terminal residue is bound to the sequencing support and was not identified

Table 2

Inhibitory activity of the fusion protein and of (desSer1 Ile29 Leu89) chicken cystatin

	(desSer1 Ile29 Leu89			
	Natural chicken cystatin		Chicken cystatin	Fusion protein
Inhibitory activity[a]	62%		40%	37%
K_i (nM) with				
Papain	0.007	0.005[b]	0.008	0.007
Cathepsin B	3.9	1.7[b]	5.7	5.6

[a] Calculated from the inhibitorily active concentration determined by titration with papain and the protein concentration measured by A_{280}

[b] Data from [1]

for stability and to clarify the interaction mechanism.

Only very recently cysteine proteinase inhibitor genes were produced by recombinant DNA methods, namely human cystatin C [33,34], stefin A [35], rat cystatin alpha [36] and stefin B [13,37]. These recombinant inhibitors will facilitate the elucidation of reaction mechanisms, of structure and function relationship and encourage studies on therapeutic effectiveness of these proteins.

Acknowledgements: The authors are grateful to S. Sokal and A. Schubert for skillful technical assistance, S. Modrow for computer help, A. Esterl for amino acid sequence analysis, W. Gebhard and G. Zuckriegl for oligonucleotide synthesis and B. Laber and U. Thiele for repurified papain and chicken cystatin and T. Popovic for cathepsin B. We also thank H. Fritz, W. Machleidt, V. Turk, M. Dolinar and W. Müller-Esterl for helpful discussions. This work has been supported by the Sonderforschungsbereich 207 der Universität München (grants G-1, H-1 and H-4) and by the International Bureau of the Kernforschungsanlage Jülich GmbH.

REFERENCES

[1] Barrett, A., Rawlings, N., Davies, M., Machleidt, W., Salvesen, G. and Turk, V. (1986) in: Proteinase Inhibitors (Barrett, A. and Salvesen, G. eds) pp.515–569, Elsevier, Amsterdam.

[2] Fossum, K. and Whitaker, J.R. (1968) Arch. Biochem. Biophys. 125, 367–375.

[3] Turk, V., Brzin, J., Longer, M., Ritonja, A., Eropkin, M., Borchart, U and Machleidt, W. (1983) Biol. Chem. Hoppe-Seyler 364, 1487–1496.

[4] Anastasi, A., Brown, M.A., Kembhavi, A.A., Nicklin, M.J.H., Sayers, C.A., Sunter, D. and Barrett, A.J. (1983) Biochem. J. 211, 129–138.

[5] Thiele, U. (1986) Diplomarbeit, Universität München, FRG.

[6] Nicklin, M.J.H. and Barrett, A.J. (1984) Biochem. J. 223, 245–253.

[7] Schwabe, C., Anastasi, A., Crow, H., McDonald, J.K. and Barrett, A.J. (1984) Biochem. J. 217, 813–817.

[8] Abrahamson, M., Ritonja, A., Brown, M.A., Grubb, A., Machleidt, W. and Barrett, A.J. (1987) J. Biol. Chem. 262, 9688–9694.

[9] Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J. and Turk, V. (1988) EMBO J. 7, 2593–2599.

[10] Beaucage, S. and Caruthers, M. (1981) Tetrahedron Lett. 22, 11859–11862.

[11] Sinha, H.D., Biernat, S. and Köster, M. (1983) Tetrahedron Lett. 24, 5843–5846.

[12] Auerswald, E.A., Schröder, W. and Kotick, M. (1987) Biol. Chem. Hoppe-Seyler 368, 1413–1425.

[13] Thiele, U., Auerswald, E.A., Gebhard, W., Assfalg-Machleidt, I., Popovic, T. and Machleidt, W. (1988) Biol. Chem. Hoppe-Seyler 369, 1167–1178.

[14] Hanahan, D. (1983) J. Mol. Biol. 166, 557–580.

[15] Birnboim, C. and Doly, J. (1979) Nucleic Acids Res. 7, 1513–1523.

[16] Heinrich, P. (1986) Guidelines for Quick and Simple Plasmid Sequencing, Boehringer, Mannheim.
[17] Strebel, K., Beck, E., Strohmaier, K. and Schaller, H. (1986) J. Virol. 57, 983–991.
[18] Remaut, E., Stanssens, P. and Fiers, W. (1981) Gene 15, 81–83.
[19] Auerswald, E.A., Hörlein, D., Reinhardt, G., Schröder, W. and Schnabel, E. (1988) Biol. Chem. Hoppe-Seyler 369, Suppl. 27–35.
[20] Kos, J., Dolinar, M. and Turk, V. (1986) in: Cysteine Proteinases and their Inhibitors (Turk, V. ed.) pp.583–591, Walter de Gruyter, Berlin.
[21] Dittmann, B., Steger, A., Wimmer, R. and Fritz, H. (1981) Biol. Chem. Hoppe-Seyler 362, 919–927.
[22] Müller-Esterl, W. (1986) in: Methods of Enzymatic Analysis IX (Bergmeyer, H.U. ed.) pp.304–316, Verlag Chemie, Weinheim.
[23] Laemmli, U.K. (1970) Nature 227, 680–685.
[24] Kyhse-Andersen, J. (1984) J. Biochem. Biophys. Methods 10, 203–209.
[25] Ogata, K., Arakawa, M., Kasahara, T., Shiori-Nakano, K. and Hiraoka, K. (1983) J. Immunol. Methods 65, 75–82.
[26] Machleidt, W., Borchart, U. and Ritonja, A. (1986) in: Advanced Methods in Protein Microsequence Analysis (Wittmann-Liebold, B. et al. eds) pp.91–107, Springer, Berlin.
[27] Sluyterman, L.A.E. and Wijdenes, J. (1970) Biochim. Biophys. Acta 200, 593–595.
[28] Machleidt, W., Machleidt, I., Müller-Esterl, W., Brzin, J., Popovic, T. and Turk, V. (1986) in: Cysteine Proteinases and their Inhibitors (Turk, V. ed.) pp.705–717, Walter de Gruyter, Berlin.
[29] Vogel, R., Assfalg-Machleidt, I., Esterl, A., Machleidt, W. and Müller-Esterl, W. (1988) J. Biol. Chem. 263, 12661–12669.
[30] Grosjeans, H. and Fiers, W. (1982) Gene 18, 192–209.
[31] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395.
[32] Schroeder, W.A., Shelton, J.B. and Shelton, J.R. (1969) Arch. Biochem. Biophys. 130, 551–557.
[33] Strauß, M., Bartsch, F.O., Stollwerk, J., Trstenjak, M., Böhning, A., Gassen, H.G., Machleidt, W. and Turk, V. (1988) Biol. Chem. Hoppe-Seyler 369, suppl. 209–218.
[34] Abrahamson, M., Dalboge, H., Olafsson, I., Carlsen, S. and Grubb, A. (1988) FEBS Lett. 236, 14–18.
[35] Strauß, M., Stollwerk, J., Lenarcic, B., Turk, V., Jany, K.H. and Gassen, H.G. (1988) Biol. Chem. Hoppe-Seyler 369, 1019–1030.
[36] Katunuma, N., Yamato, M., Kominami, E. and Ike, Y. (1988) FEBS Lett. 238, 116–118.
[37] Jerala, R., Trstenjak, M., Lenarcic, B. and Turk, V. (1988) FEBS Lett. 239, 41–48.