Evidence for the repetitive domain structure of pig calpastatin as demonstrated by cloning of complementary DNA

Emiko Takano*, Masatoshi Maki+, Masakazu Hatanaka+, Hirotaka Mori, Koichi Zenita, Toshiko Sakihama°, Reiji Kannagi, Thomas Marti†, Koiti Titani† and Takashi Murachi

Department of Clinical Science and Laboratory Medicine, Faculty of Medicine, +Institute for Virus Research, Kyoto University, Kyoto 606, Japan and *†Department of Biochemistry, University of Washington, Seattle, WA 98195, USA*

Received 19 September 1986

A clone of complementary DNA (cDNA) coding for pig heart calpastatin was isolated using synthetic oligonucleotide probes. The amino acid sequence deduced from the nucleotide sequence revealed the occurrence of a repetitive sequence at the interval of 140 amino acids, substantiating the multidomain structure of calpastatin. A portion of the sequence of 251 amino acid residues predicted for pig heart calpastatin (107 kDa) was found to be identical with that of a peptide fragment derived from pig erythrocyte calpastatin (68 kDa) and sequenced by Edman degradation.

cDNA *Calpastatin* *Domain structure*

1. INTRODUCTION

Calpastatin, the inhibitor protein acting specifically on calpain (EC 3.4.22.17; Ca^{2+}-dependent cysteine proteinase), is known to be widely distributed in mammalian and avian cells [1,2]. Previously we reported the purification of two different molecular species of calpastatin from pig heart muscle (107 kDa) and pig erythrocytes (68 kDa) [3]. Both calpastatins showed close similarity in amino acid compositions per mol and in isoelectric points. They had immunological cross-reactivity. One molecule of the 107 kDa species could bind approx. 8 calpain molecules, whereas the 68 kDa inhibitor was able to bind approx. 5 calpain molecules. These findings suggested the similarity in protein structures of the 107 and 68 kDa calpastatins, each being composed of a repetitive sequence with unit inhibitor domain. To clarify the multidomain structure of calpastatin and precursor-product relationship of the 107 and 68 kDa forms, amino acid sequence analysis of calpastatins and molecular cloning of calpastatin cDNA were attempted.

2. MATERIALS AND METHODS

2.1. *Amino acid sequence analysis*

Calpastatins were purified from pig heart muscle and pig erythrocytes as described [3]. After carboxymethylation of the purified proteins, fragmentation was performed by either CNBr treatment, endoproteinase Lys-C digestion or trypsin digestion, and the peptide fragments were separated by HPLC as in [4]. Automated Edman degradation was performed with an Applied Biosystems 470 protein sequencer using a program adapted from Hunkapiller et al. [5]. Phenylthio-

* Present address: Central Clinical Laboratory, National Kyoto Hospital, Kyoto 612, Japan
° Department of Parasitology and Immunology, Yamanashi Medical College, Yamanashi 409-38, Japan

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

hydantoin derivatives were identified in a semi-quantitative manner by two HPLC systems [6,7].

2.2. *Preparation of synthetic oligonucleotides*

Oligodeoxyribonucleotides were synthesized by the phosphoramidite method using an automatic DNA synthesizer (Applied Biosystems, model 380B) [8]. Deoxyinosine (I) was used as an alternative nucleotide for the third position of the codon with four possible combinations [9]. Oligonucleotides were labelled by phosphorylation with [γ-^{32}P]ATP and T_4 polynucleotide kinase.

2.3. *Isolation of cDNA clone*

RNA was isolated from pig heart muscle by guanidine isothiocyanate-CsCl centrifugation [10] followed by oligo(dT)-cellulose chromatography [11]. The poly(A)-containing RNA was further size-fractionated by sucrose density gradient centrifugation as in [12]. Fractions of RNA larger than 18 S were used for the construction of the cDNA library according to Okayama and Berg [13] using pcDV1 vector primer and the pL1 linker (Pharmacia). *E. coli* DH5 was transformed and resultant ampicillin-resistant colonies were screened essentially as described by Hanahan and Meselson [14] using synthetic oligonucleotide probes. The hybridization was performed at 50°C for 16 h in 6 × SET [15], 5 × Denhardt's solution [15], containing 0.5% SDS, 0.05% sodium pyrophosphate, 25 μg/ml tRNA, and ^{32}P-labelled oligonucleotides. Filters were washed in 6 × SSC [15], containing 0.05% sodium pyrophosphate, at 37°C for 10 min four times, and then once at 50 or

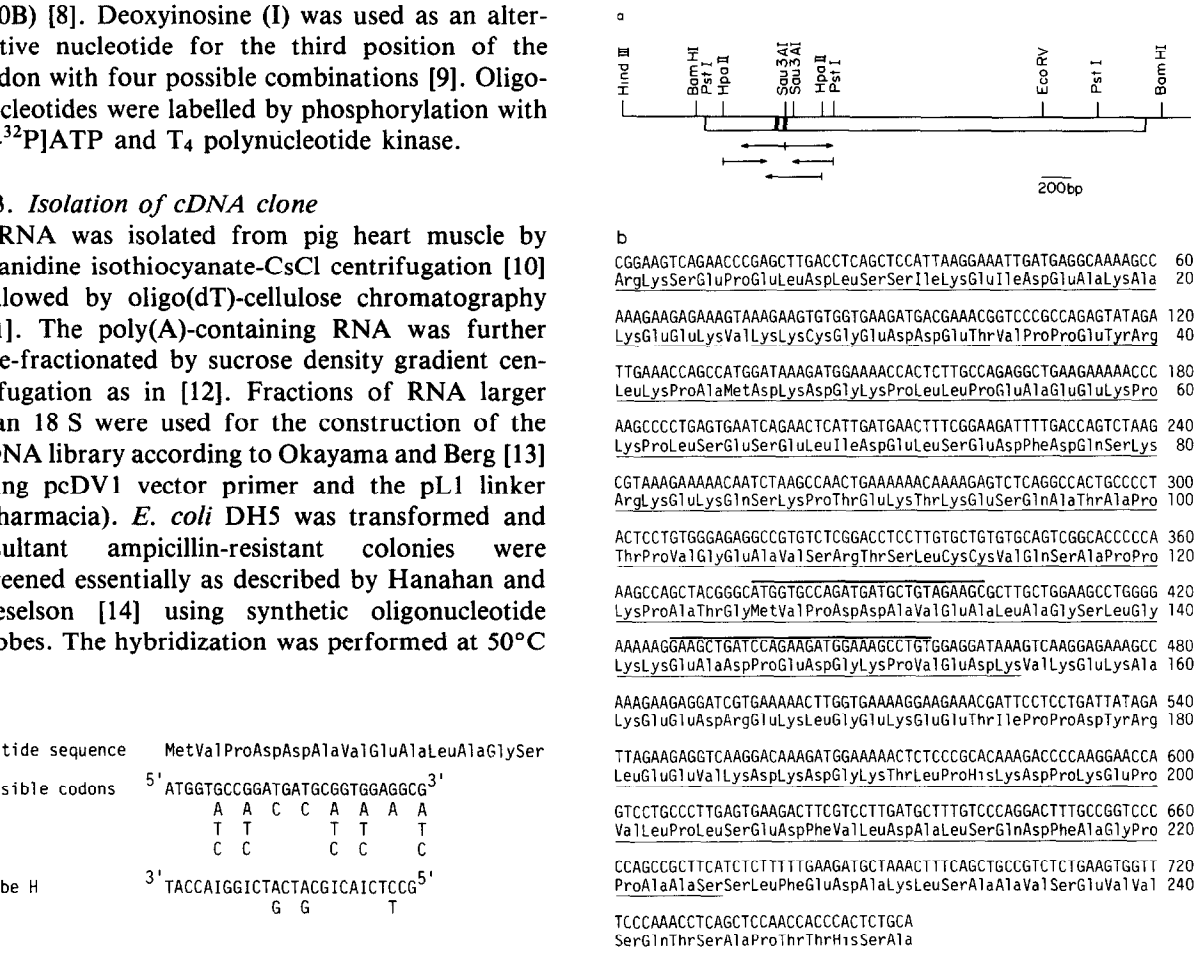

Fig.1. Synthetic oligonucleotides employed for screening cDNA libraries. Mixed oligonucleotides of 26-mer (probe H) and 29-mer (probe E) are complementary to all possible coding sequences corresponding to the peptides from heart calpastatin (A) and erythrocyte calpastatin (B), respectively. I, deoxyinosine.

Fig.2. Partial nucleotide sequence and predicted amino acid sequence of the cloned calpastatin cDNA. (a) Restriction endonuclease cleavage map of pOBCS1 and strategy for nucleotide sequencing. An open box and thin lines indicate the cDNA insert and vector DNA, respectively. Horizontal arrows indicate regions and directions of sequencing. Closed boxes indicate the positions of oligonucleotide probes. (b) Nucleotide sequence and predicted amino acid sequence. Numbering of nucleotides and amino acids is tentative. Amino acid sequences determined by Edman degradation are underlined. Overlines indicate the regions for oligonucleotide probes.

55°C. The filters were then dried and autoradiographed with intesifying screens at −70°C.

2.4. *Nucleotide sequencing*

Nucleotide sequence analysis was carried out according to Sanger et al. [16] using M13mp10 and M13mp11 for the preparation of single-stranded DNA [17]. The chain termination reaction was performed using an M13 sequencing kit (Takara Shuzo).

3. RESULTS

A pig heart cDNA library was screened by hybridization with 8 mixtures of 26-mer oligonucleotides, probe H, corresponding to the amino acid sequence of a peptide derived from a pig heart calpastatin (107 kDa) (fig.1). Out of approx. 100 000 colonies, one clone was obtained which reproducibly hybridized to the probe H. The Southern blot hybridization of the plasmid isolated from this clone, pOBCS1, gave positive signals not only to probe H but also to probe E which corresponded to the amino acid sequence of a peptide derived from pig erythrocytes (68 kDa) (fig.1). The smaller *Pst*I fragment (0.85 kb) positive to both probes was isolated and the nucleotide sequence was determined after subcloning several restriction fragments into M13mp10 and M13mp11 (fig.2a).

The predicted sequence of 251 amino acid residues was found to contain the amino acid sequences of peptide fragments which were determined by Edman degradation.

4. DISCUSSION

The present study showed that the amino acid sequence of a peptide fragment derived from pig erythrocyte calpastatin (68 kDa) (fig.1) was identical with a portion of the amino acid sequence of heart calpastatin (107 kDa) (fig.2b, amino acid residues 143–155). This result has further substantiated our previous observations that the two species of calpastatin contain common structure [3]. The finding of a long repetitive sequence at an interval of 140 amino acid residues (fig.3) also agrees well with our previous proposals that the functional unit of approx. 14 kDa comprises the multimeric structure of erythrocyte calpastatin

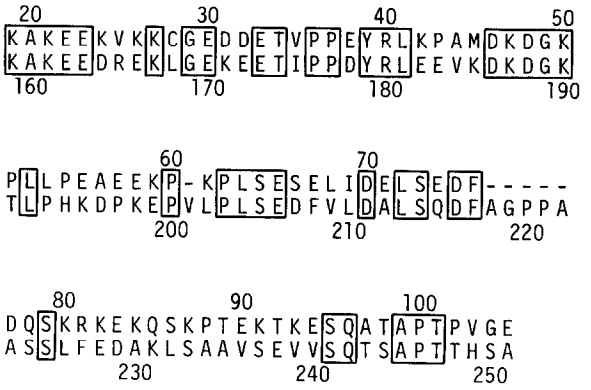

Fig.3. Internal homology of the amino acid sequence. The one-letter amino acid notation is used. Amino acid residues are placed to give the maximum homology. Boxes and dashed lines indicate identical residues and gaps, respectively.

(5-mer, 68 kDa) and heart muscle calpastatin (8-mer, 107 kDa) [3].

Comparison of the amino acid sequence of calpastatin (Asp-Lys-Asp-Gly: amino acid residues 46–49 and 186–189, fig.2b) with other known sequences revealed some homology with that of the calcium-binding domain of calmodulin [18] and other calcium-related proteins [19]. However, the obvious E-F hand structure [19] was not found. Involvement of Ca^{2+} in the activity of calpastatin remains to be seen.

Based on the inhibitory spectrum of proteinases, calpastatin, an inhibitor specific for calpain, has been shown to be distinct from a group of the cystatin family [20,21]. This may be further supported by the result of the present study that the consensus sequence of the cystatin family [20,21], Gln-Val-Val-Ala-Gly, was not found at least in the region so far determined.

The calpastatin cDNA clone isolated here contained an approx. 3.5 kb cDNA insert. However, Northern blot hybridization analysis of the size of the pig heart calpastatin mRNA revealed that the clone did not contain a full-length sequence (approx. 6 kb, not shown).

The isolated cDNA clone should serve as a useful tool for studies on the regulation of calpastatin gene expression, as well as for the isolation of full-length cDNA and subsequent nucleotide sequencing to determine the entire structure of calpastatin.

ACKNOWLEDGEMENTS

We thank Dr I. Kubota for amino acid analysis, and Miss N. Shimada for the preparation of this manuscript. This work was supported in part by Grants-in-Aid for Scientific and Cancer Research from the Ministry of Education, Science and Culture, Japan, a grant from Muscle Dystrophy Association, USA, NIH grant GM-15731 (to Dr Kenneth A. Walsh) and by a scholarship from the Swiss National Science Foundation (to T.M.).

REFERENCES

[1] Murachi, T., Tanaka, K., Hatanaka, M. and Murakami, T. (1981) Adv. Enzyme Regul. 19, 407–424.
[2] Murachi, T. (1983) in: Calcium and Cell Functions (Cheung, Y.W. ed.) vol. 4, pp. 377–410, Academic Press, New York.
[3] Takano, E., Kitahara, E., Sasaki, T., Kannagi, R. and Murachi, T. (1986) Biochem. J. 235, 97–102.
[4] Sakihama, T., Kakidani, H., Zenita, K., Yumoto, N., Kikuchi, T., Sasaki, T., Kannagi, R., Nakanishi, S., Ohmori, M., Takio, K., Titani, K. and Murachi, T. (1985) Proc. Natl. Acad. Sci. USA 85, 6075–6079.
[5] Hunkapiller, M.W., Hewick, R.M., Dreyer, W.J. and Hood, L.E. (1983) Methods Enzymol. 91, 399–413.
[6] Hunkapiller, M.W. and Hood, L.E. (1983) Methods Enzymol. 91, 486–493.
[7] Ericsson, L.H., Wade, R.D., Gagnon, J., MacDonald, R.M., Granberg, R.R. and Walsh, K.A. (1977) in: Solid Phase Methods in Protein Sequence Analysis (Previero, A. and Coletti-Previero, M.A. eds) pp. 137–142, Elsevier/North-Holland, Amsterdam, New York.
[8] Beaucage, S.L. and Caruthers, M.H. (1981) Tetrahedron Lett. 22, 1859–1862.
[9] Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, T. and Matsubara, K. (1985) J. Biol. Chem. 260, 2605–2608.
[10] Chirgwin, J.M., Przkyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[11] Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408–1412.
[12] Vamvakopoulos, N.C. and Kourides, T.A. (1979) Proc. Natl. Acad. Sci. USA 76, 3809–3813.
[13] Okayama, H. and Berg, P. (1983) Mol. Cell. Biol. 3, 280–289.
[14] Hanahan, D. and Meselson, M. (1983) Methods Enzymol. 100, 333–342.
[15] Maniatis, T., Firtisch, E.F. and Sambrook, J. (1982) in: Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[16] Sanger, F., Nicklen, S. and Coulson, A. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[17] Messing, J. (1983) Methods Enzymol. 101, 20–78.
[18] Watterson, D.M., Sharief, F. and Vanaman, T.C. (1980) J. Biol. Chem. 255, 962–975.
[19] Tufty, R.M. and Kretsinger, R.H. (1975) Science 187, 167–169.
[20] Ohkubo, I., Kurachi, K., Takasawa, T., Shiokawa, H. and Sasaki, M. (1984) Biochemistry 23, 5691–5697.
[21] Salvesen, G., Parkes, C., Abrahamson, M., Grubb, A. and Barrett, A.J. (1986) Biochem. J. 234, 429–434.