Cloning, structure and expression of cDNA for mouse contrapsin and a related protein

Kumiko OHKUBO,* Shigenori OGATA,* Yoshio MISUMI,* Noboru TAKAMI,* Hyogo SINOHARA† and Yukio IKEHARA*‡

*Department of Biochemistry, Joint Laboratory of Pathological Biochemistry, and Radioisotope Laboratory, Fukuoka University School of Medicine, Jonan-ku, Fukuoka 814-01, Japan, and †Department of Biochemistry, Kinki University School of Medicine, Sayama, Osaka 589, Japan

A cDNA clone (λMC-2) for contrapsin, a serine-proteinase inhibitor, was isolated from a λZAP mouse liver cDNA library. The 1.6 kb cDNA insert of λMC-2 contained an open reading frame that encodes a 418-residue polypeptide (46970 Da), in which a signal peptide of 21 residues was identified by comparison with the *N*-terminal sequence of the purified protein. The predicted structure (MC-2) also contained other peptide sequences determined by Edman degradation. Four potential sites for *N*-linked glycosylation were found in the molecule, accounting for the difference in molecular mass between the predicted form and the purified protein (63 kDa). Further screening of the cDNA library with an *Eco*RI–*Eco*RI fragment (510 bp) of λMC-2 as a probe yielded another cDNA clone (λMC-7), which encodes a 418-residue polypeptide (MC-7) with a calculated mass of 47010 Da. MC-2 showed 83% similarity at the amino acid level to MC-7, in contrast with 44% similarity to α_1-proteinase inhibitor. The possible reactive site (P_1–P_1') for serine proteinase is suggested to be Lys-Ala for MC-2 and Ser-Arg for MC-7. Northern-blot analysis revealed that both MC-2 and MC-7 mRNAs have the same size of 1.8 kb and are markedly induced in response to acute inflammation. Construction of the expression plasmids pSVMC-2 and pSVMC-7 and their transfection into COS-1 cells demonstrated that pSVMC-2 directs the synthesis of a 63 kDa form whereas pSVMC-7 expresses a 56 kDa form. The difference in molecular mass between the two may be explained by the fact that the MC-7 sequence contains three potential sites for *N*-glycosylation, one site less than that of MC-2.

INTRODUCTION

Human plasma contains several inhibitors of proteolytic enzymes that together form about 10% of its protein content [1]. Of the inhibitors, α_1-proteinase inhibitor (α_1-PI; also called α_1-antitrypsin) has been extensively studied [2], since its variant forms are closely related to the occurrence of severe diseases, including pulmonary emphysema [2,3] and a fatal bleeding disorder [4,5]. α_1-PI is now known as the archetype of a superfamily called 'serpin', a group of serine-proteinase inhibitors with similar sequences [6]. The serpin superfamily includes α_1-PI, α_1-antichymotrypsin, antithrombin III, $C\bar{1}$ inhibitor etc. The specificity of each of these inhibitors is primarily dependent on a single amino acid residue at its reactive site (P_1).

In contrast with the human serpins, those of animals have not been so extensively studied. It has been shown that mouse plasma contains two major trypsin inhibitors, one of which corresponds to α_1-PI and the other is contrapsin [7]. Although α_1-PI will inhibit most of serine proteinases, contrapsin was found to exert no inhibitory activity against elastase, chymotrypsin or thrombin [7,8]. Subsequently Hill *et al.* [9] reported a partial sequence (215 amino acid residues) of contrapsin deduced from the cDNA sequence, demonstrating that contrapsin is highly similar to human α_1-antichymotrypsin rather than to α_1-PI. Thus it is suggested that the genes for mouse contrapsin and human α_1-antichymotrypsin are the descendants of a single gene that have evolved since rodent and primate divergence to encode proteins with different functions.

In the present study we established the complete sequence of contrapsin by cloning and sequencing its cDNA, which was confirmed by comparison with chemically determined sequences of the purified protein. We also demonstrate the presence of another protein having 83% similarity to contrapsin, although its inhibitory activity remains to be determined.

MATERIALS AND METHODS

Materials

[^{35}S]Methionine (> 800 Ci/mmol) and [α-^{32}P]dCTP (400 or 3000 Ci/mmol) were obtained from Du Pont–New England Nuclear (Boston, MA, U.S.A.). λZAP bacteriophage vector and the Gigapack Gold kit were obtained from Stratagene (San Diego, CA, U.S.A.), the cDNA synthesis kit was from Amersham Corp. (Arlington Heights, IL, U.S.A.) and the Sequenase DNA-sequencing kit was from United States Biochemical Corp. (Cleveland, OH, U.S.A.). Various DNA-modifying enzymes and restriction endonucleases were from Nippon Gene (Toyama, Japan), Takara Shuzo (Kyoto, Japan) and New England Biolabs (Boston, MA, U.S.A.).

Purification of contrapsin

Contrapsin was purified from mouse serum by a method essentially the same as that used for purification of α_1-PI [10]. In brief, mouse serum (20 ml) was subjected to Affi-Gel Blue chromatography, $(NH_4)_2SO_4$ precipitation (40–80% saturation), DE-52 DEAE-cellulose chromatography (twice, firstly at pH 8.8 and secondly at pH 5.5) and Sephadex G-100 chromatography.

Abbreviation used: α_1-PI, α_1-proteinase inhibitor.

‡ To whom correspondence should be addressed.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X55147 and X55148.

During these steps, contrapsin was detected by measuring its trypsin-inhibiting activity [11] in combination with SDS/PAGE (10% gels) [12].

Isolation of contrapsin-derived peptides

Purified contrapsin (1 mg) was digested with lysyl endopeptidase (enzyme/substrate ratio by wt, 1:200) in 0.3 ml of 50 mM-Tris/HCl buffer, pH 9.0, at 30 °C for 8 h [13]. The resulting peptide fragments were freeze-dried and then subjected to reverse-phase h.p.l.c. on a TSKgel ODS-120T column (0.46 cm × 25 cm) with a linear gradient from 8% to 70% (v/v) acetonitrile in 0.1% (v/v) trifluoroacetic acid [14]. Fractions of major peptide peaks were collected, freeze-dried and used for amino acid sequencing.

Amino acid sequencing

The purified contrapsin (about 40 μg) and its derived peptides (about 7 μg of each) were sequenced on an Applied Biosystems model 477A gas-phase sequencer with an on-line model 120A phenylthiohydantoin derivative analyser as described elsewhere [14,15].

cDNA library construction and screening

Polyadenylated RNA was prepared from BALB/c mouse liver [16,17] and used for construction of the following cDNA library. Double-stranded cDNA was prepared according to the method established by Gubler & Hoffman [18]. *Eco*RI-linkers were ligated to the cDNA and digested with *Eco*RI. Linked cDNA was purified and ligated to λZAP phage vector [19]. The ligated cDNA was packaged into bacteriophages by using the Gigapack Gold kit followed by plating on to dishes with BB4 bacterial hosts for amplification as the cDNA library stock [20]. A part of the library stock (4×10^5 clones) was screened with ^{32}P-labelled cDNA (1.6 kb) prepared from λCPi-21, a cDNA clone for rat contrapsin-like proteinase inhibitor (K. Ohkubo, Y. Misumi, S. Ogata, N. Takami & Y. Ikehara, unpublished work). A clone (λMC-2) with the longest cDNA insert (1.6 kb) was isolated. An *Eco*RI–*Eco*RI fragment (510 bp) was prepared from the MC-2 cDNA insert, oligolabelled and used as a probe for further screening of the cDNA library (2×10^5 clones). Positive clones obtained were subjected to the second screening. The resulting single positive phages were cultured with helper phage R408, and insert cDNAs were obtained in pBluescriptSK$^-$ plasmid vector by the automatic excision process [19]. The cDNA inserts were characterized by restriction-endonuclease mapping [17]. We finally obtained two similar cDNA clones with inserts of 1.6 kb, λMC-2 and λMC-7.

DNA sequencing

The entire inserts and restriction fragments of cDNAs were subcloned into pUC118 or pUC119. After single-stranded DNAs were isolated with the aid of helper bacteriophages (M13K07), both strands of all regions were sequenced by the dideoxynucleotide chain-termination method [21] with the use of the Sequenase DNA-sequencing kit.

Northern-blot analysis

Turpentine (0.5 ml/100 g body wt.) was injected intramuscularly into male BALB/c mice for induction of acute inflammation. At the indicated times after the injection, livers were excised, frozen in liquid N_2 and then subjected to RNA extraction [16]. The extracted RNA (5 μg of each) was electrophoresed in 1.5% agarose gels containing 6% formaldehyde. The separated RNA was transferred to a Durapore membrane (Millipore) and hybridized at 42 °C for 20 h with ^{32}P-labelled *Eco*RI–*Eco*RI fragment (510 bp) prepared from λMC-2 [14]. The RNA transferred to the membrane was also hybridized with either of the following synthetic oligonucleotides: 5′-GCTGGTAATATGGCCTT-3′ (for MC-2) and 5′-GTTTGTAATCTTCTAGA-3′ (for MC-7). The membranes were washed, dried and autoradiographed [14,17].

Construction of expression plasmids

cDNA inserts of λMC-2 and λMC-7 were prepared by digestion of each plasmid with appropriate restriction enzymes, purified by electrophoresis and then inserted into the corresponding sites of pSVM expression vector [22]. Each insert orientation was confirmed by restriction-endonuclease mapping. The expression plasmids thus prepared are designated as pSVMC-2 and pSVMC-7.

Transfection and analysis of expressed proteins

Each plasmid (10 μg) was transfected into 3×10^6 COS-1 cells as described previously [23,24]. The transfected cells were cultured in Dulbecco's modified Earle's medium (Nissui Seiyaku, Tokyo, Japan) containing 10% (v/v) fetal-calf serum in 10 cm dishes for 2 days. The cells were incubated at 37 °C for 4 h with [^{35}S]-methionine (100 μCi/dish) in 5 ml of methionine-free Eagle's minimum essential medium (Nissui Seiyaku). Each medium separated from cells was used for immunoprecipitation with anti-contrapsin IgG [7,24]. Samples of the medium and the immunoprecipitates were analysed by SDS/PAGE (10% gels) followed by fluorography. Apparent molecular masses were determined by co-electrophoresis of ^{35}S-labelled marker proteins [25]: transferrin (78 kDa), albumin (66 kDa), α_1-PI (56 kDa) and haptoglobin β-chain (36 kDa).

RESULTS AND DISCUSSION

Purification and amino acid sequencing of contrapsin

Contrapsin was purified from mouse serum by a procedure slightly different from that used in the previous study [7]. The purified protein was found to be a single component with an

Table 1. *N*-Terminal sequence of purified mouse contrapsin

	Component 1		Component 2	
Cycle no.	Amino acid	Amount (pmol)	Amino acid	Amount (pmol)
1	Phe	275	Asp	66
2	Pro	180	Gly	96
3	Asp	148	Thr	60
4	Gly	204	Lys	94
5	Thr	125	Glu	81
6	Lys	198	Met	114
7	Glu	127	Asp	43
8	Met	225	Ile	107
9	Asp	98	Val	92
10	Ile	210	Phe	67
11	Val	184	His	15
12	Phe	152	Glu	46
13	His	42	Xaa	
14	Glu	89	Gln	33
15	His	17	Asp	28
16	Gln	81		
17	Asp	69		
18	Xaa			
19	Gly	60		
20	Thr	23		

Table 2. Amino acid sequences of contrapsin-derived peptides

Four major peptides (about 7 μg of each), obtained by lysyl endopeptidase digestion and h.p.l.c., were subjected to amino acid sequencing. The amino acid sequences determined are assigned to those predicted by the MC-2 cDNA sequence (see Fig. 2).

	Peptide 1		Peptide 2		Peptide 3		Peptide 4	
Cycle no.	Residue	Amount (pmol)	Residue	Amount (pmol)	Residue	Amount (pmol)	Residue	Amount (pmol)
1	Asn	125	Thr	108	Ile	294	Phe	230
2	Pro	207	Arg	96	Ser	112	Ser	105
3	Asp	240	Ala	245	Phe	178	Ile	216
4	Thr	132	Leu	192	Asp	170	Ala	223
5	Asn	76	Tyr	170	Pro	134	Ser	78
6	Ile	126	Gln	165	Gln	115	Asn	62
7	Val	130	Thr	77	Asp	86	Tyr	96
8	Phe	98	Glu	108	Thr	65	Arg	84
9	Ser	59	Ala	131	Phe	93	Leu	87
10	Pro	87	Phe	109	Glu	87	Glu	68
11	Leu	96	Thr	51	Xaa		Glu	85
12	Ser	31	Ala	86	Glu	66	Asp	57
13	Ile	75	Asp	93	Phe	52	Val	69
14	Ser	25	Phe	75	Tyr	58	Leu	48
15	Ala	82	Gln	61	Leu	79	Pro	40
16	Ala	68	Gln	76	Asp	32	Glu	35
17	Leu	34	Pro	52	Glu	28	Met	39
18			Thr	28	Lys	14	Gly	28
19			Glu	33			Ile	24
20			Ala	54			Lys	9
21			Lys	12				
Predicted positions...		46–62		135–155		197–214		294–313

apparent molecular mass of 63 kDa when analysed by SDS/PAGE and immunochemical techniques (results not shown; see ref. [7]). The protein was analysed for the *N*-terminal sequence (Table 1). Sequencing data demonstrated the presence of two components in the sample that have different *N*-termini: a major component (component 1 in Table 1) that starts with phenylalanine, and a minor one (component 2) that starts with aspartic acid. The sequence of component 2, however, is essentially the same as that of component 1, lacking only the first two residues of the component 1. Thus the results indicate that contrapsin primarily has an *N*-terminal sequence starting with phenylalanine. The difference of two residues will cause no significant difference in molecular mass between the two components when analysed by SDS/PAGE [7]. The presence of component 2 suggests that the *N*-terminal two residues of contrapsin, in part, could be post-translationally cleaved off, for example, by the ectoenzyme dipeptidyl peptidase IV. Peptides/proteins having the Phe-Pro sequence at the *N*-terminus serve as good substrates for the enzyme [15,26].

The purified contrapsin was also subjected to cleavage with lysyl endopeptidase, and the resulting peptides were separated by reverse-phase h.p.l.c. Four major peptides obtained were sequenced (Table 2).

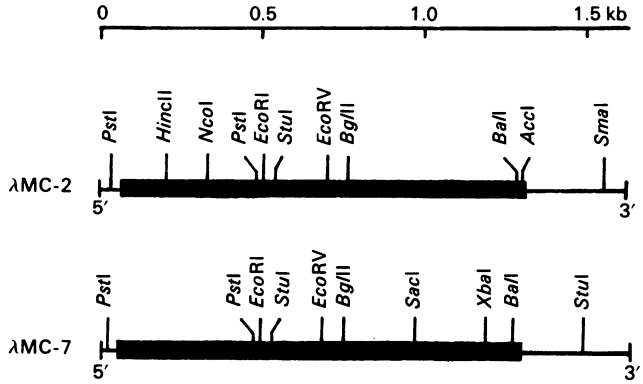

Fig. 1. Restriction maps of cDNAs for mouse contrapsin and a related protein

The entire cDNA inserts of λMC-2 and λMC-7 and fragments prepared by appropriate restriction-enzyme digestions were subcloned into plasmid vector pUC118 or pUC119, and sequenced as described in the Materials and methods section. The protein coding region in each cDNA is indicated by a thick closed bar.

Cloning and sequencing of cDNA clones

Screening of the mouse liver cDNA library with the ^{32}P-labelled cDNA of rat CPi-21 yielded 20 positive clones, of which λMC-2 had the longest insert of 1.6 kb and was analysed by restriction mapping, as shown in Fig. 1. The nucleotide sequence of the cDNA insert was determined (Fig. 2). The sequence contains an open reading frame that encodes a 418-amino acid-residue polypeptide with a calculated molecular mass of 46970 Da. The first 21 residues (−21 to −1) containing a hydrophobic amino acid cluster may represent a signal peptide, since the sequence is followed by the *N*-terminal sequence determined for the purified protein (Table 1). The predicted sequence also contains all the other peptide sequences determined by Edman degradation. Lys-Ala at positions 363–364 may be the reactive site (P_1–P_1') for trypsin, as suggested by Hill *et al.* [9]. The same reactive site Lys-Ala was also reported for a trypsin inhibitor of bovine pancreas [27,28]. Four potential sites (Asn-Xaa-Ser/Thr) for *N*-linked glycosylation are found in the molecule, possibly accounting for the difference in molecular mass between the predicted precursor and the mature form (63 kDa). The sequence previously reported by Hill *et al.* [9] corresponds to

```
                                    5'---- CAACCCTGAACATCAGGAGTCAGCTATCACAGAGGCTCTGCAGCCTGGAGAACAGAGAAG      60
                                           ***
ATG GCC TTC ATT GTA GCT ATG GGG ATG ATC TTA ATG GCT GGA ATC TGT CCT GCT GTC CTA TGC TTC CCA GAT GGC     135
Met Ala Phe Ile Val Ala Met Gly Met Ile Leu Met Ala Gly Ile Cys Pro Ala Val Leu Cys Phe Pro Asp Gly       4
-21                                                                                          +1
ACA AAG GAA ATG GAC ATT GTA TTC CAT GAA CAC CAA GAC AAT GGG ACA CAA GAT GAC AGT CTC ACA TTG GCC TCC     210
Thr Lys Glu Met Asp Ile Val Phe His Glu His Gln Asp Asn Gly Thr Gln Asp Asp Ser Leu Thr Leu Ala Ser      29

GTC AAC ACT GAC TTT GCC TTC AGC CTG TAC AAG AAG CTG GCT TTG AAG AAT CCA GAT ACA AAT ATT GTC TTC TCC     285
Val Asn Thr Asp Phe Ala Phe Ser Leu Tyr Lys Lys Leu Ala Leu Lys Asn Pro Asp Thr Asn Ile Val Phe Ser      54

CCA CTT AGC ATC TCA GCT GCC TTG CGC CTT GTG TCC CTG GGA GCA AAG GGC AAG ACC ATG GAA GAG ATT CTA GAA     360
Pro Leu Ser Ile Ser Ala Ala Leu Arg Leu Val Ser Leu Gly Ala Lys Gly Lys Thr Met Glu Glu Ile Leu Glu      79

GGC CTC AAG TTC AAT CTC ACA GAG ACC CCT GAA GCA GAC ATC CAC CAG GGC TTT GGC AAC CTC CTA CAG AGT CTC     435
Gly Leu Lys Phe Asn Leu Thr Glu Thr Pro Glu Ala Asp Ile His Gln Gly Phe Gly Asn Leu Leu Gln Ser Leu     104

AGC CAG CCA GAA GAC CAG GAT CAG ATA AAC ATA GGC AAT GCC ATG TTT ATT GAA AAG GAC CTG CAG ATC CTG GCA     510
Ser Gln Pro Glu Asp Gln Asp Gln Ile Asn Ile Gly Asn Ala Met Phe Ile Glu Lys Asp Leu Gln Ile Leu Ala     129

GAA TTC CAT GAG AAG ACA AGG GCT CTG TAC CAG ACT GAG GCC TTC ACA GCA GAC TTC CAG CAG CCT ACT GAG GCC     585
Glu Phe His Glu Lys Thr Arg Ala Leu Tyr Gln Thr Glu Ala Phe Thr Ala Asp Phe Gln Gln Pro Thr Glu Ala     154

AAA AAC CTC ATC AAT GAC TAT GTG AGC AAT CAG ACC CAG GGG ATG ATC AAG GAA CTC ATC TCA GAA CTG GAT GAG     660
Lys Asn Leu Ile Asn Asp Tyr Val Ser Asn Gln Thr Gln Gly Met Ile Lys Glu Leu Ile Ser Glu Leu Asp Glu     179

AGG ACA TTG ATG GTG CTG GTG AAT TAC ATC TAC TTT AAA GGC AAA TGG AAG ATA TCC TTT GAC CCC CAG GAC ACA     735
Arg Thr Leu Met Val Leu Val Asn Tyr Ile Tyr Phe Lys Gly Lys Trp Lys Ile Ser Phe Asp Pro Gln Asp Thr     204

TTT GAG TCT GAG TTC TAC TTG GAT GAG AAG AGA TCT GTG AAG GTT CCC ATG ATG AAA ATG AAG TTA CTG ACC ACA     810
Phe Glu Ser Glu Phe Tyr Leu Asp Glu Lys Arg Ser Val Lys Val Pro Met Met Lys Met Lys Leu Leu Thr Thr     229

CGC CAC TTC CGT GAT GAG GAG CTA TCG TGC TCT GTG TTG GAG CTG AAG TAC ACA GGA AAT GCC AGC GCC CTG CTC     885
Arg His Phe Arg Asp Glu Glu Leu Ser Cys Ser Val Leu Glu Leu Lys Tyr Thr Gly Asn Ala Ser Ala Leu Leu     254

ATC CTC CCT GAC CAG GGC AGG ATG CAG CAG GTG GAA GCC AGC TTA CAA CCA GAG ACC CTG AGG AAA TGG AGG AAA     960
Ile Leu Pro Asp Gln Gly Arg Met Gln Gln Val Glu Ala Ser Leu Gln Pro Glu Thr Leu Arg Lys Trp Arg Lys     279

ACT TTG TTT CCC AGC CAA ATA GAG GAG CTA AAC CTG CCC AAG TTC TCC ATC GCT AGT AAC TAC AGG CTG GAG GAG    1035
Thr Leu Phe Pro Ser Gln Ile Glu Glu Leu Asn Leu Pro Lys Phe Ser Ile Ala Ser Asn Tyr Arg Leu Glu Glu     304

GAT GTC CTT CCA GAA ATG GGG ATT AAG GAA GTC TTC ACC GAA CAA GCT GAC CTA TCT GGG ATC ACA GAA ACC AAG    1110
Asp Val Leu Pro Glu Met Gly Ile Lys Glu Val Phe Thr Glu Gln Ala Asp Leu Ser Gly Ile Thr Glu Thr Lys     329

AAA CTG AGT GTG TCT CAG GTG GTC CAC AAG GCT GTG CTG GAT GTG GCT GAG ACA GGC ACA GAA GCA GCT GCT GCC    1185
Lys Leu Ser Val Ser Gln Val Val His Lys Ala Val Leu Asp Val Ala Glu Thr Gly Thr Glu Ala Ala Ala Ala     354

ACA GGG GTT ATT GGT GGC ATT CGT AAG GCC ATA TTA CCA GCT GTG CAT TTC AAC AGG CCA TTC CTG TTT GTT ATC    1260
Thr Gly Val Ile Gly Gly Ile Arg Lys Ala Ile Leu Pro Ala Val His Phe Asn Arg Pro Phe Leu Phe Val Ile     379

TAT CAC ACA AGT GCT CAG AGT ATC CTC TTT ATG GCC AAA GTC AAT AAC CCC AAG TAG ACCTAAAGCTCACCAAGTTCTCA    1340
Tyr His Thr Ser Ala Gln Ser Ile Leu Phe Met Ala Lys Val Asn Asn Pro Lys ***                              397

TGGTTCTGTCCAGGGCTCTGGGACAAAGTCTGCATGTGGATCTCTATATACATCCTGTCATTCAAGCTCTGATTGGCTGTTGCAAAGTTGGCTTAGACA    1439

GCTACATTAACTATCTCTGTGATCCCACATGCACATAAGACTTTGGACTCTCAGTGTCAGGCTCTCGGTTGTCTTGGGAGCATCTGTGCATATTTCTGA    1538

ACCTGGAATCTGCCTTTATTCTTCTTCCCGGGTGACTCCTCTTTATGTGTCTGTACCCCAAACCTAGGCATTTGCAAATACACAGAGTTCCAAT        1632
```

Fig. 2. Nucleotide sequence of the cDNA and deduced amino acid sequence of mouse contrapsin

Both nucleotide and predicted amino acid residues are numbered on the right. Amino acid residues −21 to −1 comprise a putative signal peptide. Amino acid residues that have been determined by protein sequence analysis of purified contrapsin are underlined. Circled asparagine residues represent potential *N*-linked glycosylation sites. A double-underline indicates the reactive site (P_1–P'_1) of the inhibitor. The stop codons limiting the open reading frame are indicated by asterisks.

a *C*-terminal half of the present sequence (positions 183–397). The two sequences are essentially the same except for two positions: we found methionine (rather than valine) at position 183 and threonine (rather than isoleucine) at position 326.

The 5′-end *Eco*RI–*Eco*RI fragment (510 bp) of λMC-2 was prepared and used as a probe for further screening of the cDNA library. Another cDNA clone (λMC-7) with a similar insert (1.6 kb) was obtained (Fig. 1). The cDNA insert of λMC-7 was also sequenced for prediction of its primary structure. Fig. 3 shows an amino acid sequence (MC-7) predicted by the λMC-7 cDNA sequence in comparison with that of contrapsin (MC-2). MC-7 has the same number of amino acid residues (418 residues) as those of MC-2, showing 83.1 % similarity to the latter (88.1 % similarity at the nucleotide level). The alignment of the two sequences suggests that MC-7 has a signal peptide with one residue less than that of MC-2, and that its mature form contains an additional residue as compared with that of the latter. MC-7 has three potential *N*-linked glycosylation sites, one site less than those of MC-2. Its possible reactive site for serine proteinase is suggested to be Ser-Arg, in contrast with Lys-Ala in contrapsin MC-2. The two sequences of MC-2 and MC-7 are also compared with those of mouse α_1-PI [29], rat α_1-PI [24,30], human α_1-antichymotrypsin [31,32] and human protein C inhibitor [33]. Both MC-2 and MC-7 are found to have a significantly higher degree of similarity (50–55 %) to human α_1-antichymotrypsin than to mouse α_1-PI (42–44 %), rat α_1-PI (42–43 %) and human protein C inhibitor (42 %). It may be noted that the reactive region (P_5–P'_5) of human protein C inhibitor (IFTFRSARLN) is

```
MC-2      -21 MAFIVAMGMILMAGICPAVLCFPDGTKEMDIVFHEHQDNGTQDDSLTLASVNTDFAFSLYKKLALKNPDTNIVFSPLSISA   60
MC-7      -20 ----A-L-*--------T----S-D-WGI--LL-KN-ES--P--------I----------E-V-----K----------   60

MC-2    ALRLVSLGAKGKTMEEILEGLKFNLTETPEADIHQGFGNLLQSLSQPEDQDQINIGNAMFIEKDLQILAEFHEKTRALYQTEAFTADFQQ  150
MC-7    --A--------N-L----------S---------H---R-----------------------------A----------K  150

MC-2    PTEAKNLINDYVSNQTQGMIKELISELDERTLMVLVNYIYFKGKWKISFDPQDTFESEFYLDEKRSVKVPMMKMKLLTTRHFRDEELSCS  240
MC-7    ----------TD-----------------------------------------------F----------  240

MC-2    VLELKYTGNASALLILPDQGRMQQVEASLQPETLRKWRKTLFPSQIEELNLPKFSIASNYRLEEDVLPEMGIKEVFTEQADLSGITETKK  330
MC-7    ----------F----------------------W-S-KTRK-G--Y------STD-N-K*-I---L----I-SK--------G--D  329

MC-2    LSVSQVVHKAVLDVAETGTEAAAATGVIGGIRKAILPA**VHFNRPFLFVIYHTSAQSILFMAKVNNPK   397
MC-7    ----------F-F-F-SRR-QTMT-Q------M--S--GV-TT------T---   398
```

Fig. 3. Comparison of the amino acid sequence of contrapsin (MC-2) with that predicted for MC-7

The amino acid residues are shown in the single-letter code. Amino acid residues of MC-7 that are identical with those of MC-2 are indicated by a dash. An arrow and double arrowheads show a signal-peptide cleavage site and a possible reactive site (P_1–P_1') for serine proteinase respectively. Dots indicate potential *N*-linked glycosylation sites. Asterisks indicate gaps introduced into the sequences so that they could be aligned.

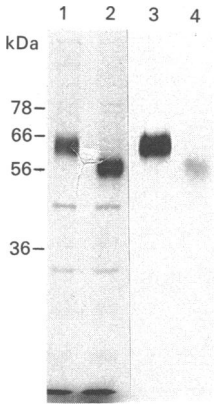

Fig. 4. Analysis of expression-plasmid-directed products

COS-1 cells transfected with pSVMC-2 (lanes 1 and 3) or pSVMC-7 (lanes 2 and 4) were incubated with [^{35}S]methionine at 37 °C for 4 h. Culture media were separated from the cells. Samples of each medium were directly subjected to SDS/PAGE (10% gels), followed by fluorography (lanes 1 and 2). The remainders of each sample were subjected to immunoprecipitation with anti-contrapsin IgG, and the resulting immunoprecipitates (lanes 3 and 4) were analysed as above.

quite similar to the sequence IFGFRSRRLN (positions 357–366) of MC-7. This fact suggests the possibility that the reactive site (P_1–P_1') of MC-7 is Arg-Ser at positions 361–362, instead of Ser-Arg (362–363). Taken together, these results indicate that MC-7 is an additional member of the serpin family, although its inhibitory activity remains to be determined.

Transfection and analysis of expressed proteins

The expression plasmids pSVMC-2 (for contrapsin MC-2) and pSVMC-7 (for MC-7) were constructed from pSVM [22] and the corresponding cDNAs, and each plasmid was transfected into COS-1 cells. Labelling experiments demonstrated that the transfected cells synthesized and secreted single major proteins (Fig. 4). The protein expressed by transfection with pSVMC-2 (Fig. 4, lane 1) had a molecular mass of about 63 kDa, identical with that of the protein purified from mouse serum, and was immunoprecipitated with anti-contrapsin IgG (lane 3). The product directed by pSVMC-7 (Fig. 4, lane 2) had 56 kDa, and only a trace amount of the product could be immunoprecipitated with anti-contrapsin IgG (lane 4). Thus the secreted forms of MC-2 and MC-7 are clearly different from each other in molecular mass and immunochemical properties, although the sequences predicted for their precursors show a high degree of similarity (83%) and almost the same molecular mass (47 kDa) for the two. The difference in molecular mass between the two products may be mostly due to a difference in their sugar moieties, since MC-2 has four *N*-glycosylation sites whereas MC-7 has three sites (Fig. 3).

Northern-blot analysis

Fig. 5 shows the Northern-blot hybridization analysis, for which a ^{32}P-labelled *Eco*RI–*Eco*RI fragment (510 bp) prepared from λMC-2 was used as a probe. A single band of 1.8 kb was identified in the total RNA isolated from control mouse liver (Fig. 5, lane 1). To examine the response of the contrapsin mRNA concentration to acute inflammation, we prepared liver RNAs from mice at the indicated times after turpentine injection. It was found that 1.8 kb mRNA is markedly increased at 12 h after the injection (Fig. 5, lane 2) and then decreased to or below the control value by 86 h (lanes 3 and 4). The results are in good agreement with those previously obtained with a translation system *in vitro* [34], confirming that contrapsin is an acute-phase reactant [34,35]. This is in contrast with the finding that mouse α_1-PI is not induced by acute inflammation [35,36].

For specific detection of each mRNA for MC-2 and MC-7, we prepared two 17-mer oligonucleotides corresponding to coding regions of their reactive sites in which the sequences are more dissimilar than other parts. It was confirmed that the synthetic probe for MC-2 hybridizes to the 1.8 kb RNA with the same response to acute inflammation as was observed with the *Eco*RI–*Eco*RI fragment. In addition, the probe for MC-7 hybridized to 1.8 kb RNA that was indistinguishable from that for MC-2 (results not shown). The RNA concentration detected

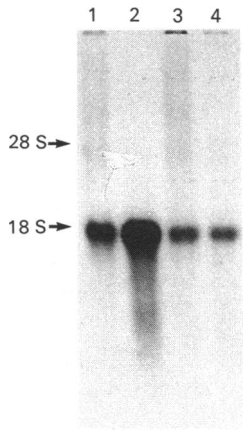

Fig. 5. Northern-blot analysis of mouse liver contrapsin mRNA

Total liver RNA was prepared from mice at 0, 12, 48 and 86 h (lanes 1–4 respectively) after turpentine injection. RNA (5 μg of each) was electrophoresed, transferred to a membrane and hybridized with ^{32}P-labelled *Eco*RI–*Eco*RI fragment (510 bp) of λMC-2, followed by autoradiography. RNA size markers are 28 S rRNA and 18 S rRNA of rat liver.

showed a similar time course to that described above in response to acute inflammation. These results suggest that the mRNA of MC-7 is also 1.8 kb long and induced by acute inflammation similarly to contrapsin MC-2.

We thank Dr. Y. Sakaki (Kyushu University) for preparation of oligonucleotides and Y. Hayashi for assistance in purification of mouse contrapsin. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Travis, J. & Salvesen, G. S. (1983) Annu. Rev. Biochem. **52**, 655–709
2. Carrell, R. W., Jeppsson, J.-O., Laurell, C.-B., Brennan, S. O., Owen, M. C., Vaughan, L. & Boswell, D. R. (1982) Nature (London) **298**, 329–334
3. Laurell, C.-B. & Ericksson, S. (1963) Scand. J. Clin. Lab. Invest. **15**, 132–140
4. Lewis, J. H., Iammarino, R. M., Spero, J. A. & Hasiba, H. (1978) Blood **51**, 129–137
5. Owen, M. C., Brennan, S. O., Lewis, J. H. & Carrell, R. W. (1983) N. Engl. J. Med. **309**, 694–698
6. Carrell, R. & Travis, J. (1985) Trends Biochem. Sci. **10**, 20–24
7. Takahara, H. & Sinohara, H. (1982) J. Biol. Chem. **257**, 2438–2446
8. Takahara, H. & Sinohara, H. (1983) J. Biochem. (Tokyo) **93**, 1411–1419
9. Hill, R. E., Shaw, P. H., Boyd, P. A., Baumann, H. & Hastie, N. D. (1984) Nature (London) **311**, 175–177
10. Ikehara, Y., Miyasato, M., Ogata, S. & Oda, K. (1981) Eur. J. Biochem. **115**, 253–260
11. Dietz, A. A., Rubinstein, H. M. & Hodges, L. (1974) Clin. Chem. **20**, 396–399
12. Laemmli, U. K. (1979) Nature (London) **227**, 680–685
13. Tsunasawa, T., Sugihara, A., Masaki, T., Sakiyama, F., Takeda, Y., Miwatani, T. & Narita, K. (1987) J. Biochem. (Tokyo) **101**, 111–121
14. Misumi, Y., Ogata, S., Hirose, S. & Ikehara, Y. (1990) J. Biol. Chem. **265**, 2178–2183
15. Ogata, S., Misumi, Y. & Ikehara, Y. (1989) J. Biol. Chem. **264**, 3596–3601
16. Chirwin, J. M., Prybyla, E., MacDonald, R. J. & Rutter, W. J. (1979) Biochemistry **18**, 5294–5299
17. Misumi, Y., Tashiro, K., Hattori, M., Sakaki, Y. & Ikehara, Y. (1988) Biochem. J. **249**, 661–668
18. Gubler, U. & Hoffman, B. J. (1983) Gene **25**, 263–269
19. Short, J. M., Fernandez, J. M., Sorge, J. A. & Huse, W. D. (1988) Nucleic Acids Res. **16**, 7583–7600
20. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, pp. 368–369, Cold Spring Harbor Laboratory, Cold Spring Harbor
21. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
22. Misumi, Y., Ohkubo, K., Sohda, M., Takami, N., Oda, K. & Ikehara, Y. (1990) Biochem. Biophys. Res. Commun. **171**, 236–242
23. Oda, K., Takami, N., Fujiwara, T., Misumi, Y. & Ikehara, Y. (1989) Biochem. Biophys. Res. Commun. **163**, 194–200
24. Misumi, Y., Sohda, M., Ohkubo, K., Takami, N., Oda, K. & Ikehara, Y. (1990) J. Biochem. (Tokyo) **108**, 230–234
25. Misumi, Y., Misumi, Y., Miki, K., Takatsuki, A., Tamura, G. & Ikehara, Y. (1986) J. Biol. Chem. **261**, 11398–11403
26. McDonald, J. K. & Schwabe, C. (1977) in Proteinases in Mammalian Cells and Tissues (Barrett, A. J., ed.), vol. 2, pp. 311–391, North-Holland, Amsterdam
27. Kassell, B., Radicevic, M., Ansfield, M. J. & Laskowski, M., Jr. (1965) Biochem. Biophys. Res. Commun. **18**, 255–258
28. Anderson, S. & Kingston, I. B. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 6838–6842
29. Sifers, R. N., Ledley, F. D., Reed-Fourquet, L., Ledbetter, D. H., Ledbetter, S. A. & Woo, S. L. C. (1990) Genomics **6**, 100–104
30. Chao, S., Chai, K. X., Chao, L. & Chao, J. (1990) Biochemistry **29**, 323–329
31. Chandra, T., Stackhouse, R., Kidd, V., Robson, J. & Woo, S. (1983) Biochemistry **22**, 5055–5061
32. Morii, M. & Travis, J. (1983) J. Biol. Chem. **258**, 12749–12752
33. Suzuki, K., Deyashiki, Y., Nishioka, J., Kurachi, K., Akira, M., Yamamoto, S. & Hashimoto, S. (1987) J. Biol. Chem. **262**, 611–615
34. Yamamoto, K., Suzuki, Y. & Sinohara, H. (1988) Biochem. Int. **16**, 921–928
35. Suzuki, Y., Yoshida, K., Ichiyama, T., Yamamoto, T. & Sinohara, H. (1990) J. Biochem. (Tokyo) **107**, 173–179
36. Baumann, H., Latimer, J. J. & Glibetic, M. D. (1986) Arch. Biochem. Biophys. **246**, 488–493

Received 25 September 1990/12 November 1990; accepted 15 November 1990