FEBS Letters 457 (1999) 262–270 FEBS 22492

CA-MMP: a matrix metalloproteinase with a novel cysteine array, but without the classic cysteine switch[1]

Duanqing Pei*

Department of Pharmacology, 3-249 Millard Hall, 435 Delaware St. S.E., University of Minnesota, Minneapolis, MN 55455, USA

Received 30 June 1999

Abstract A matrix metalloproteinase (MMP)-like gene was identified in mouse to contain a conserved MMP catalytic domain and an RRRR motif. It lacks a classic cysteine switch, but it possesses two novel motifs: a cysteine array (Cys-X_6-Cys-X_8-Cys-X_{10}-Cys-X_3-Cys-X_2-Cys), and a novel Ig-fold. It is named CA-MMP after the distinct cysteine array motif, and little is known about its biochemical function. In an attempt to characterize CA-MMP activity, the full-length sequence was expressed in mammalian cells and its product found to be cell-associated without detectable secretion. In light of this unusual finding, a chimera combining the catalytic domain of CA-MMP with the prodomain of stromelysin-3 was constructed to express a fully active enzyme in mammalian cells. Purified CA-MMP catalytic domain expresses proteolytic activity against protein substrates in an MMP inhibitor sensitive fashion. Taken together, it is concluded that CA-MMP is an MMP with distinct structure, biochemical properties and evolutionary history that may define a new subclass of the MMP superfamily.

Key words: ECM; MMP; Proteolysis; Latency

1. Introduction

The extracellular matrix (ECM), made of various collagens, elastins, proteoglycans, and glycoproteins, undergoes extensive remodeling or breakdown under many physiological as well as pathological conditions [1,2]. The primary enzymes responsible for ECM breakdown appear to belong to a family of metal-dependent, neutral pH optimal endopeptidases called matrix metalloproteinases (MMPs) [2–4]. Synthesized as latent zymogens, MMPs undergo obligatory activation before they can express any proteolytic activity against ECM component [5]. Once activated, MMPs can be inhibited by proteinase inhibitors both general such as α_2M and specific such as TIMPs [6,7]. Throughout evolution, various mechanisms such as zymogen latency and shields of proteinase inhibitor have been evolved gradually to fine tune MMP activities in normal tissues.

Functional repertoire of domains or motifs which confer substrate specificity and regulatory mechanism for MMPs discovered so far include: (1) a signal peptide for secretion, (2) a cysteine switch motif PRCGXPD for latency, (3) an RXKR motif for furin mediated zymogen activation, (4) a catalytic domain built around an HEXXHXXXXXH motif, (5) a fibronectin-like domain for binding to substrates, (6) a hemopexin-like domain for both substrate and inhibitor interactions, (7) a transmembrane domain for cellular localization [4,6]. It seems that the ancestral MMP has a domain structure resembling that of matrilysin containing the signal peptide, PRCGXPD and the catalytic domain [8]. Based on this design, additional acquisition of various domains and motifs led to the current family of about 20 MMPs. For example, gelatinase A is made of all the above domains except RXKR and transmembrane domains, whereas its activators, the MT-MMPs, possess all except the fibronectin-like motif [9–11]. This modular design reflects both the general principle of gene evolution as well as specific need of matrix degrading enzymes. While all these motifs are integral parts of MMPs, only two are unique for this gene family: the catalytic and the PRCGXPD domains.

The symbiotic relationship between the catalytic and PRCGXPD domains is illustrated elegantly by the cysteine switch model proposed by Van Wart and Birkedal-Hanson [12]. The Cys residue in the PRCGXPD motif is proposed to bind to a catalytic zinc coordinated by the HEXXHXXGXXH motif in the active site, blocking the zinc from interacting with an H_2O molecule, thus keeping the enzyme in latent form until activation with the removal of the prodomain [12]. A critical prediction of this model is that MMPs can be activated by perturbing this cys-zinc interaction by chemical, physical as well as proteolytic interactions. Indeed, oxidants which react with the thiol group of cysteine, chaotropic agents such as organomercurial compounds which relax the prodomain or proteinases such as trypsin which attack the prodomain directly, can all induce MMPs to undergo autoactivation to generate fully active MMP enzyme [13–16]. Although the cysteine switch model governs the latency/activation of all known MMPs, it is not clear whether any alternative mechanism has evolved to keep similar MMP catalytic domain latent.

Here I report the cloning of CA-MMP: a novel mouse MMP-like gene lacking a cysteine switch in the prodomain. Furthermore, the discovery of CA-MMP brings two novel functional domains to the repertoire of MMP structures: a distinct cysteine array and a novel Ig-fold domain. Characterization of its products confirms that CA-MMP is a matrix metalloproteinase expressing strong gelatinolytic activity sensitive to both synthetic as well as natural MMP inhibitors. During the preparation of this paper, two groups have reported the sequence of the human counterpart of CA-MMP. Velasco et al. have described the sequence, tissue distribution and chromosome localization of MMP23, while Gururajan et al. reported the same sequence and its duplicated copy as

*Corresponding author. Fax: +1 (612) 625-8408.
E-mail: peixx003@tc.umn.edu

[1] The nucleotide sequence reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number AF085742.

PII: S0014-5793(99)01046-7

MMP21/22 near the CDC2L1-2 locus at 1p36.3 [17–19]. Recognized as a novel MMP gene with distinct structure in both reports, very little has been described about the enzymic function of human MMP21/22/23 and its key structural features [19].

2. Materials and methods

2.1. Cell lines and reagents

General chemical reagents were either from Fisher (Pittsburg, PA, USA) or Sigma (St. Louis, MO, USA). MDCK cells and COS 7 were obtained and maintained as described previously [20–22]. DNA restriction and modification enzymes were purchased from Promega (Madison, WI, USA). Oligonucleotide primers were synthesized at the University of Minnesota microchemical core facility. In general, COS cells are used for transient gene expression largely because the pCR3.1 expression vector system is very efficient in COS cells due to the presence of SV40 T antigen. MDCK is the choice of cells for stable expression of recombinant protein as described [21]. Anti-FLAG M2 monoclonal antibody, its agarose beads, and FLAG peptide were purchased from Kodak (Rochester, NY, USA). Human recombinant TIMP2 was obtained from Dr. R. Fridman (Wayne State University, MI, USA).

2.2. Cloning and analysis of CA-MMP cDNA and genomic locus

The bulk of the mouse CA-MMP gene is derived from an EST clone AA238244 by direct sequencing. The missing 5′ of CA-MMP is recovered by cloning the genomic locus of CA-MMP from screening the genomic library of strain SVJ129 using CA-MMP cDNA as probe. The entire coding region of CA-MMP was mapped to the genomic locus by PCR and sequencing as shown in Fig. 1, lower panel. A composite clone was assembled by PCR using high fidelity *pfu* enzyme and confirmed by sequencing. Conceptual translation of the cDNA sequence was carried out via internet at http://www.expasy.ch/. Sequence alignment was performed via Internet using program Multalin version 5.3.3 at http://www.expasy.ch/www/tools.html using blosum62 with Gap weight: 12; Gap length: 2.

2.3. Tissue distribution of CA-MMP by RT-PCR

Premade mouse cDNA panels (Clontech, Palo Alto, CA, USA) were amplified with two primers: 5′ TCA ATT CCT CAC TCG GAC CC 3′, 5′ CTC TAT GGC TGC CTG GAC 3′ at the following conditions: 2 min at 94°C for denaturation, 35 cycles of 10 s at 94°C, 30 s at 54°C for annealing and 30 s at 72°C for extension, followed by 10 min extension at 72°C to generate a 420 bp fragment. To control for the amount of cDNA used in each reaction, a parallel amplification using primers designed from the house-keeping gene GAPDH was performed under the same experimental conditions [20].

2.4. Construction of ST3/CA-MMP-LYG and CA-MMP expression vectors

The full-length expression vector was constructed by amplifying the entire open reading frame of CA-MMP with primers: CA-MMP 5′ CAT CCC TTT CCC ATT CCG; CA-MMP-FLAG 3′ GTC ACT TGT CAT CGT CGT CCT TGT AGT CAT TCC TCA CTC GGA CCC, and then inserting the PCR product into the *Eco*RV site of pCR3.1uni as described. The resulting construct contains a FLAG tag at its 3′ end. The chimera ST3/CA-MMP-LYG was constructed by high fidelity PCR to fuse the prodomain of stromelysin-3 $(ST3)_{1-97}$ with the catalytic domain of $CA\text{-}MMP_{80-255}$ as follows [23]. Briefly, the prodomain of ST3 was isolated by high fidelity PCR using two primers: ST3 5′ CGG ATG GCT CCG GCC GCC TGG CTC CGC AGC, and ST3 3′ CCT CTT CTG TCG GTT GCG GGC, and the catalytic domain of CA-MMP using primers: CA-MMP-cat 5′ CGA CAG AAG AGG TAC ACG CTG ACA CCG GCC AGG, and CA-MMP-cat-FLAG 3′ GTC ACT TGT CAT CGT CGT CCT TGT AGT CGC CAT AGA GTC GGT GTA AC. The resulting two PCR fragments were combined and re-amplified with the ST3 5′ and CA-MMP-cat-FLAG 3′ primers to give rise to the chimera fragment. The chimera product was subsequently cloned into expression vector pCR3.1uni at the *Eco*RV site. Clones with the correct orientation were identified by restriction digestion and confirmed by DNA sequencing.

2.5. DNA transfection and immunofluorescence staining

The expression construct pCR3.1uni-CA-MMP-FLAG was transfected into COS cells for efficient expression and the resulting products were analyzed by immunoprecipitation using M2 monoclonal antibody as described [20,21]. For stable expression, the full-length or the chimera construct was transfected into MDCK cells separately and stable clones were selected, isolated and characterized as described previously [20,21]. Positive clones were identified by Western blot using anti-FLAG M2 monoclonal antibody and expanded as described [20,21]. For immunofluorescence, MDCK cells stably transfected with control or full-length CA-MMP were fixed, stained with M2 monoclonal antibody as the first antibody (1:200 dilution), FITC conjugated rabbit anti-mouse secondary antibodies (Calbiochem, CA, USA) and visualized under a Nikon fluorescence microscope as described [24].

2.6. Purification and characterization of CA-MMP catalytic domain

Since the prodomain of ST3 contains a furin recognition site for MMP activation [23], the chimera protein is expected to be processed by furin into a fully active form, i.e. the catalytic domain. To purify the secreted catalytic domain, a stable clone with high level of expression was expanded in 150 mm dishes to confluence. The cultures were then washed with phosphate-buffered saline (PBS) twice, replenished with serum free DMEM media and allowed to incubate for 48 h before harvesting. A total of approximate 500 ml of conditioned media was collected, cleared by centrifugation, and fractionated by affinity chromatography using the M2 antibody resin (Kodak, NY, USA) as described [21]. The proteolytic activity of purified CA-MMP was analyzed against denatured type I collagen in the presence or absence of MMP inhibitors BB94 or TIMP2 as described [22].

3. Results

3.1. Cloning of mouse CA-MMP cDNA and characterization of its exon/intron structure

During a search for novel MMP related sequences in the public EST database maintained in the National Center for Biotechnology Information, several EST hits were analyzed further to identify MMP-like genes from the mouse genome. One of such ESTs, AA238244, reveals extensive homology to human stromelysin-3 gene with identity reaching $>40\%$ at the catalytic domain. Subsequent sequencing of the entire EST clone, however, predicted a putative MMP with novel domains at the C-terminus (Fig. 1). To complete the open reading frame, the genomic locus of this gene was isolated from mouse genomic library (SVJ129 strain) which covers a 19 kb segment, where approximately 2.5 kb encodes the entire exon 8 of this particular gene as determined by direct sequencing and PCR amplification (Fig. 1, lower panel). This is the smallest MMP-like gene identified from any mammalian genome.

3.2. Structural features of CA-MMP

As shown in Fig. 1, a conceptual translation of the open reading frame indicates that this novel MMP-like gene encodes: (1) a hydrophobic segment which could be the signal peptide, (2) a potential furin recognition motif: RRRR, (3) a conserved MMP catalytic domain, (4) a novel carboxyl terminus and (5) three potential N-glycosylation sites (N93, N149, N233) (Fig. 1).

The putative prodomain of CA-MMP lacks a cysteine switch, but contains a furin processing site [12,23]. Although all MMPs discovered so far contain a well conserved PRCGXPD motif for latency, CA-MMP appears to lack such a motif. The three cysteine residues at the N-terminus do not appear to be available for latency because they are either upstream or part of a hydrophobic region (bold letters

```
gccgcccatccctttcccattccgccccagttttgcgatccaggtccctacgtcgcgcgg
ccccttgtggccctaacccgattggagtctgccatgggctgccgggcctgtctccgtccg
                                        M  G  C  R  A  C  L  R  P    9
gaggcgtccggcgcggtacagggccgctggttgggggctgccctgagcgggttgtgcctt
 E  A  S  G  A  V  Q  G  R  W  L  G  A  A  L  S  G  L  C  L   29
ctctccgcgttggcgttgctggagtggctaggggcgccgaccgagaccgcctggagggca
 L  S  A  L  A  L  L  E  W  L  G  A  P  T  E  T  A  W  R  A   49
gctcagggaaatgtagatgcgccaaatgtgggcagctccactgcccaggtacccagactg
 A  Q  G  N  V  D  A  P  N  V  G  S  S  T  A  Q  V  P  R  L   69
ttgaccatgtcggtaacccgaagacgccgctacacgctgacaccggccaggctgcgctgg
 L  T  M  S  V  T  R  R  R  R  Y  T  L  T  P  A  R  L  R  W   89
gaccacttcaacctcacatacaggagttctttccttcccccggaatcttttaagcccggaa
 D  H  F  N  L  T  Y  R  V  L  S  F  P  R  N  L  L  S  P  E  109
gagacgcgaaggggcctggctgctgcctttcgaatgtggagtgatgtttccccattcagt
 E  T  R  R  G  L  A  A  A  F  R  M  W  S  D  V  S  P  F  S  129
ttccgtgaggtggcccctgaacgtcccagtgacctcaagataggtttctacccagtcaac
 F  R  E  V  A  P  E  R  P  S  D  L  K  I  G  F  Y  P  V  N  149
cacaccgactgcttggtctctgcagtgcaccactgctttgatggtcccacaggtgaactg
 H  T  D  C  L  V  S  A  V  H  H  C  F  D  G  P  T  G  E  L  169
gcccacgctttcttcccaccccacggtggcattcactttgatgacagcgagtactgggtc
 A  H  A  F  F  P  P  H  G  G  I  H  F  D  D  S  E  Y  W  V  189
ttgggccccacacgctacagttggaagaaaggtgtttggctcacaaacctggtgcacgtg
 L  G  P  T  R  Y  S  W  K  K  G  V  W  L  T  N  L  V  H  V  209
gcagcccatgagattggccatgcactgggcctgatgcactcacagcaagatcaggcgctc
 A  A  H  E  I  G  H  A  L  G  L  M  H  S  Q  Q  D  Q  A  L  229
atgcacctcaatgccacattgcgaggctggaaggcactgtcccaggatgaactgtggggg
 M  H  L  N  A  T  L  R  G  W  K  A  L  S  Q  D  E  L  W  G  249
ttacaccgactctatggctgcctggaccggatttttgtgtgtgcatcctgggcacgaaag
 L  H  R  L  Y  G  C  L  D  R  I  F  V  C  A  S  W  A  R  K  269
ggattttgtgatgtccgccagaggctcatgaagaggctctgccccagaagctgtgatttc
 G  F  C  D  V  R  Q  R  L  M  K  R  L  C  P  R  S  C  D  F  289
tgttatgaattcccatttcccacggtggccaccaccacatcacccaccagaacgaaaaca
 C  Y  E  F  P  F  P  T  V  A  T  T  T  S  P  T  R  T  K  T  309
aggttggtgagagagggtaggaacatgaccttccactgtggacagaagatcctacacaag
 R  L  V  R  E  G  R  N  M  T  F  H  C  G  Q  K  I  L  H  K  329
aaagggaaagtatactggtacaaggaccaggaacccctggagttctcctaccctggctat
 K  G  K  V  Y  W  Y  K  D  Q  E  P  L  E  F  S  Y  P  G  Y  349
ctggccctgggtgaggcacagctgagcatcattgccaacgcagtcaacgaaggcacctac
 L  A  L  G  E  A  Q  L  S  I  I  A  N  A  V  N  E  G  T  Y  369
acatgtgtggtacgccggcaccagcgtgttctcagcacctactcctggagggtccgagtg
 T  C  V  V  R  R  H  Q  R  V  L  S  T  Y  S  W  R  V  R  V  389
aggaattgacactgctcagagcctctgtgctggaaggcagcagataaagcactttctctc
 R  N  -                                                       391
tga
```

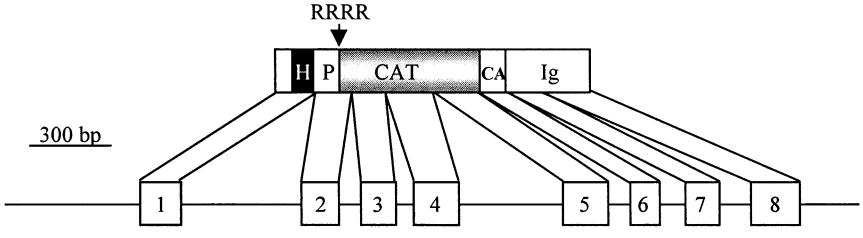
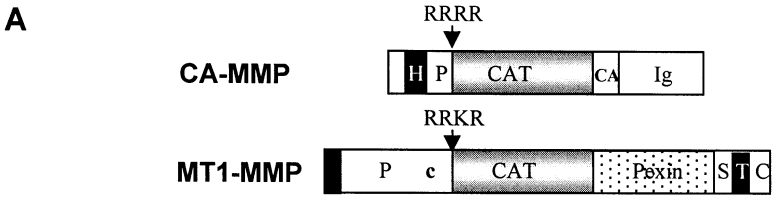


Fig. 1. Primary structure of mouse CA-MMP. Upper panel: Nucleotide and amino acid sequences of mouse CA-MMP. The numbering on the right is for the amino acid sequence. The hydrophobic segment at the N-terminus is presented in bold. The putative furin recognition motif RRRR is boxed. The zinc catalytic site is underlined with the three His and one Glu residues bolded, followed by the M turn (bold) and putative salt bridge Asp (bold). The six Cys residues which form a cysteine array are italicized, bolded and underlined. The vertical bars indicate the positions of seven introns. Lower panel: Exon/intron and domain structure of mouse CA-MMP: The motifs and domains mentioned above are presented schematically. RRRR and the downward arrow indicate the potential furin processing signal. The relationship between the cDNA structure (upper line) and the genomic locus (lower line) is illustrated with connecting lines specifying the position of each exon (boxes 1–8). H, hydrophobic segment; P, prodomain; CAT, catalytic domain; CA, cysteine array; Ig, immunoglobulin C2 type fold.

in Fig. 1, upper panel). Surprisingly, this gene contains a conserved furin recognition motif RRRR, which could serve as a signal for zymogen activation, as demonstrated for MMP11 and MMP14 [23,25,26], suggesting that the removal of the prodomain process may be required for CA-MMP activation.

The CA-MMP catalytic domain contains two novel inserts. Alignment of the catalytic domain with those of a few selected MMPs indicates strong homology surrounding the zinc binding catalytic site: HEXGHXXGXXH with the conserved M turn as well as the Asp residue for salt bridge formation with the N-terminus of the mature enzyme (Fig. 2A) [27]. Interest-



B

```
mMMP8          FLLTPGSPKWTHTNLTYRIINHTPQLSR-AEVKTAIEKAFHVWSVASPLTFTEILQG---
XCOL4          YSTVAKSSAWQKKDLTYRILNFTPDLPQ-ADVETAIQRAFKVWSDVTPLTFTRIYNE---
MT1            --YAIQGLKWQHNEITFCIQNYTPKVGE-YATYEAIRKAFRVWESATPLRFREVPYAYIR
MMP12          FREMPGGPVWRKHYITYRINNYTPDMNR-EDVDYAIRKAFQVWSNVTPLKFSKINTG---
ENAMLYSIN       YRLFPGEPKWKKNTLTYRISKYTPSMSS-VEVDKAVEMALQAWSSAVPLSFVRINSG---
MMP11          --FVLSGGRWEKTDLTYRILRFPWQLVQ-EQVRQTMAEALKVWSDVTPLTFTEVHEG---
mCA-MMP        YTLTPARLRWDHFNLTYRVLSFPRNLLSPEETRRGLAAAFRMWSDVSPFSFREVAPE---
                         * :  :*: :  .. .:     .   :  *:: *. . *: * .:

mMMP8          -----EADINIAFVSRDHGDNSP------FDGPNGILAHAFQPGQGIGGDAHFDSEETWT
XCOL4          -----VSDIEISFTAGDHKDNSP------FDGSGGILAHAFQPGNGIGGDAHFDEDETWT
MT1            EGHEKQADIMIFFAEGFHGDSTP------FDGEGGFLAHAYFPGPNIGGDTHFDSAEPWT
MMP12          -----MADILVVFARGAHGDFHA------FDGKGGILAHAFGPGSGIGGDAHFDEDEFWT
ENAMLYSIN       -----EADIMISFENGDHGDSYP------FDGPRGTLAHAFAPGEGLGGDTHFDNPEKWT
MMP11          -----RADIMIDFARYWHGDDLP------FDGPGGILAHAFFPKTHREGDVHFDYDETWT
mCA-MMP        ----RPSDLKIGFYPVNHTDCLVSAVHHCFDGPTGELAHAFFPPHG--G-IHFDDSEYWV
                     :*: : *    * *          ***  * ****: *     *  ***  * *.

mMMP8          QDS---------KNYNLFLVAAHEFGHSLGLSHSTDPGALMYPNYAYREPSTYSLPQDDI
XCOL4          KTS---------EIYNLFLVAAHEFGHSLGLSHSTDQGALMYPTYSNTDPKTFQLPQDDI
MT1            VRNEDL------NGNDIFLVAVHELGHALGLEHSSDPSAIMAPFYQWMDTENFVLPDDDR
MMP12          THS---------GGTNLFLTAVHEIGHSLGLGHSSDPKAVMFPTYKYVDINTFRLSADDI
ENAMLYSIN       MGT---------NGFNLFTVAAHEFGHALGLAHSTDPSALMYPTYKYKNPYGFHLPKDDV
MMP11          IGDD--------QGTDLLQVAAHEFGHVLGLQHTTAAKALMSAFYTFRYP--LSLSPDDC
mCA-MMP        LGPTRYSWKKGVWLTNLVHVAAHEIGHALGLMHSQQDQALMHLNATLRGWK--ALSQDEL
                          ::. .*.**:** *** *:    *:*             *. *:

mMMP8          NGIQTIYG
XCOL4          NAIQYLYG
MT1            RGIQQLYG
MMP12          RGIQSLYG
ENAMLYSIN       KGIQALYG
MMP11          RGVQHLYG
mCA-MMP        WGLHRLYG
                .:: :**
```

Fig. 2. Domain structure of CA-MMP. A: Comparison of domain structures between CA-MMP and MT1-MMP: The catalytic domain and the RxK/RR signal are aligned between CA-MMP and MT1-MMP. H, hydrophobic; P, prodomain; CA, cysteine array; Ig, immunoglobulin C2 type fold; c, cysteine switch; CAT, catalytic domain; Pexin, hemopexin-like domain; S/TM/C, stem/transmembrane/cytosolic domains. B: Alignment of MMP catalytic domains. The catalytic domains from the following representative MMPs: mouse MMP8 (mMMP8) and CA-MMP, frog collagenase 4 (XCOL4), human MT1-MMP (MT1), MMP12, Enamelysin and MMP11 are aligned to illustrate the degree of homology between CA-MMP and the selected MMPs. Homologous positions to the CA-MMP sequence are marked bold. Identical residues are marked by *, while the conserved ones by dots underneath. The two unique inserts for CA-MMP are boxed. C, cysteine array. The cysteine rich segment of CA-MMP is aligned with protein sequences deposited in the NCBI database to identify the cysteine array in every six, eight, ten, three and two residues. U41274, Z66520, U97407, U80452, Z71261, Z73896, P34269, U13072, U22380, Z79639 and U58755 are accession numbers for putative genes from the *C. elegans* genome. The consensus is depicted at the bottom with bold letters. D, type C2 Ig-fold. The sequence downstream of the cysteine array has two segments of homology with those of CD80, DCC, IL1R and CD48, similar to the type C2 Ig-fold. DCC, deleted in colon cancer; IL1R, interleukin 1 receptor.

ingly, two novel inserts are present in the catalytic domain. The first, C_{153}LVSAVHHC$_{161}$, appears to be able to form a loop via a potential disulfide bond whereas the other, T_{193}**R**YSW**KK**GV_{201}, is positively charged (Fig. 2A).

The cysteine array motif is encoded by an anciently shuffled exon. The sequence downstream of the catalytic domain does not have any specific feature other than a distinct Cys pattern of Cys-X_6-Cys-X_8-Cys-X_{10}-Cys-X_3-Cys-X_2-Cys (Fig. 2B). Analysis of CA-MMP genomic structure reveals that this motif is encoded by a single exon: exon 6, suggesting an evolutionary conserved domain (Fig. 1). Indeed, when this sequence was blasted alone against sequence deposited at NCBI, the Cys pattern matches perfectly with more than 11 entries derived from the *C. elegans* genome project with unknown function (Fig. 2B). The same basic pattern repeats four times in one particular sequence, U97407 (Fig. 2B). In light of this finding, this MMP-like gene is named CA-MMP for cysteine array matrix metalloproteinase.

The carboxyl end of CA-MMP contains a novel domain with an Ig-like C2 type fold. Downstream of the cysteine

C

```
CA-MMP 256 CLDRIFVCASWARKGFCDVRQRLMKRLCPRSCDFC 290
U41274 586 CSDRKHFCSHWKSAGFCEIFMNYMKKNCPASCGLC 621
U41274 487 CEDLRVDCLVLVSQRYCKISQNFMKSYCAKSCGFC 521
Z66520  22 CTDKHIHCFFWSQEGECEVNPRWMKKHCQKACGTC  56
U97407 502 CTDRHTQCSSWSRSGECTKNQLWMTENCRKSCNKC 536
U97407 458 CFNENQCCASWAASGECSRNTAYMNEWCKASCGVC 492
U97407 612 CVDYHRSCAGWARVGECQ-KNPWMAENCRSSCNSC 645
U97407 568 CFNENVCCAVWGLMGECRKNTRYMACNCRVSCGHC 602
U80452  71 CKDRHDLCKFWSSIGECTSNKNWMEDHCPVSCDFC 105
U80452 111 CIDRHRLCGFWATIRECETNAVWMLSNCPKACKAC 145
Z71261 754 CADYHYDCAAWARRGEC-LKNKWMPENCRRSCNTC 787
Z71261 636 CSDRHTNCAMWSRSGECNKNPLWMSENCRSSCQKC 670
Z71261 592 CFNENECCGPWSAKGECQKNPVYMNVWCKASCRQC 626
Z73896 483 CEDKNLFCSYWAKIGECNSESKFMKIFCKASCGKC 517
Z73896 394 CEDLNAHCGMWEQLGHCQHSVKYMAHYCRKACNLC 428
P34269 406 CHDRHISCAYWRGQNFCTRRRQWMAENCQATCGWC 440
P34269 363 CYNEDPCCNQWSRQNECRTNTVYMNRYCRKSCGLC 397
U13072 384 CEDRITVCWWTADRCRSPAIYQVMSSLCPKTCKFC 418
U22380 199 CEDSHSNCAAWAKANECNKNPNWMRPNCKKSCGTC 233
Z79639 255 CEDKHSLCCFWATTGECSRNPFWMKLNCAKTCGTC 289
Z79639 301 CKDYYKECSFWAVLGECQSNPFYMRPNCQKSCKSC 335
U58755 248 CRDRTNLCWRWIDRCKSFFFEQIMKEFCSLSCGYC 282
Consensus  C-D----C--W---geC------M--nC--sC--C
```

D

```
                                                    
CA-MMP 332 KVYWYKDQEPLEFSYPGYLA-------------LGEAQLSIIANAVNE-GTYTCVV
CD80    62 RIYWQKDQQMVLSIISGQVEVWPEYKNRTFPDIINNLSLMILALRLSDKGTYTCVV
DCC    270 SFTWLRGEEVIQLRSKKYSL-------------LGGSNLLISNVTDDDSGTYTCVV
IL1R    52 TITWYKDDSKTPVSTEQASRIHQHKEKLWFVP----------A-KVEDSGHYYCVV
CD48   135 DYTWYEDSGPFPQRNPGYV-------------------LEITITPHNKSTFYTCQV
```

Fig. 2 (*continued*).

array, there are 101 residues with a potential disulfide bridge (Cys^{322}-Cys^{371}). Within this bridging pair of cysteines, there is a recognizable Ig-like C2 type fold similar to those in CD80, DCC, IL1R, and CD48 (Fig. 2C) [28–31]. Thus, this motif along with the cysteine array defines a novel carboxyl terminus for CA-MMP.

3.3. Expression profile of CA-MMP in mouse embryo and adult tissues

To begin to probe the physiological function of CA-MMP, its expression profile was determined in mouse embryo and adult tissues by RT-PCR as described [20]. As shown in Fig. 3, heart, spleen and lung are the main organs/tissues expressing CA-MMP at relatively high levels, while tissues such as brain, liver, skeletal muscle, kidney and testis are basically negative. On the other hand, 7 days old mouse embryo expresses CA-MMP at the highest level when compared to day 11, 15 and 17 embryos. It appears that CA-MMP expression begins to increase on day 17 after its level drops from day 11 to 15 (Fig. 3, lanes 9–12). The onset of CA-MMP expression has not been determined due to the lack of embryos prior to day 7.

3.4. Expression of full-length CA-MMP

Based on sequence analysis, there is a hydrophobic segment at the N-terminus of CA-MMP which could be consistent with either a signal peptide or type II transmembrane domain. To begin to distinguish this possibility and explore the biochemical function of CA-MMP, a full-length CA-MMP expression vector, pCR3.1uni-CA-MMP-FLAG, was constructed to contain a FLAG tag at its carboxyl terminus (Fig. 4A). In a transient transfection assay into COS cells, both the supernatants and cell lysates were analyzed to detect any CA-MMP product using the anti-FLAG M2 antibody as

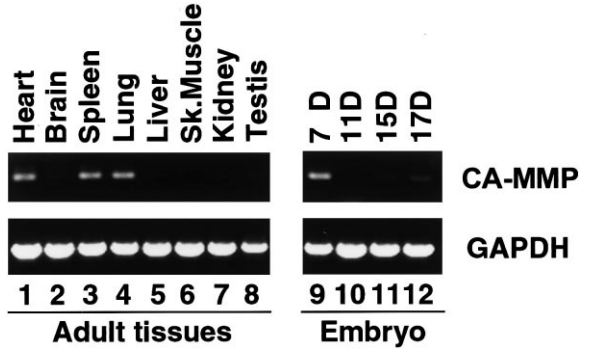


Fig. 3. Expression of CA-MMP in mouse embryo and adult tissues. Reverse transcribed cDNA (0.5 ng aliquot each) from a multiple tissue panel (Clontech, CA, USA) including heart, brain, spleen, lung, liver, skeleton muscle, kidney testis (lanes 1–8) were amplified to give rise to the 420 bp fragment of MT5-MMP. Similar cDNA preparations from whole embryos aged day 7, 11, 15, 17 (E7 to 17 in lanes 9–12) were analyzed as the adult tissues. To control for the amount of cDNA in each reaction, a parallel PCR reaction with GAPDH specific primers was carried out as suggested by the supplier.

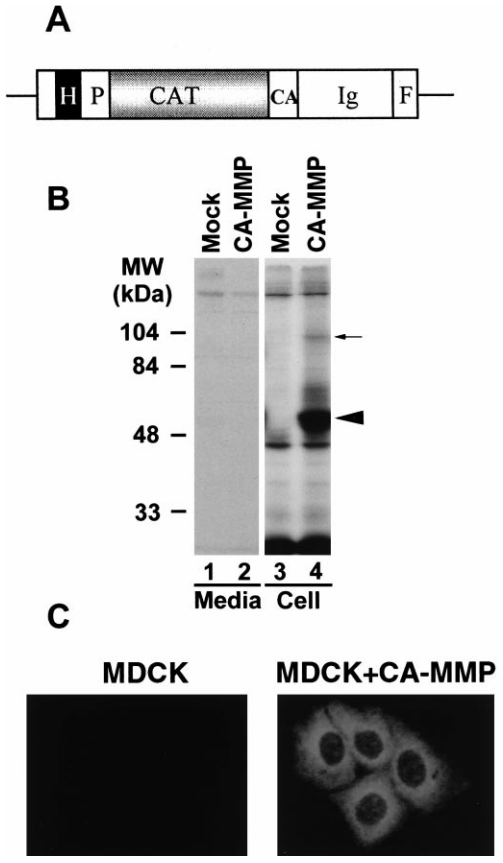


Fig. 4. CA-MMP is associated with cells. A: CA-MMP-FLAG expression construct. The full-length CA-MMP was fused in-frame with a FLAG epitope at its carboxyl terminus to facilitate detection and purification (see Section 2). F, FLAG epitope. B: CA-MMP is associated with cells. Control (lanes 1, 3) and mouse CA-MMP (lanes 2, 4) expression vectors were transfected into COS cells. 48 h post-transfection, the cells were labeled with [^{35}S]Met (100 μCi/ml, Amersham, IL, USA) for 3 h. The supernatants (lanes 1, 2) and cell lysates (lanes 3, 4) were immunoprecipitated with anti-FLAG M2 monoclonal antibody as described in Section 2. The arrowhead indicates the main CA-MMP species while small arrow above the arrowhead marks a putative dimer of CA-MMP. C: Immunofluorescence of CA-MMP products. MDCK cells stably transfected with control (left panel) or full-length CA-MMP (right panel) vectors were stained with M2 monoclonal antibody and then FITC conjugated secondary antibody as described [24].

described [21]. As shown in Fig. 4B, a main CA-MMP product at 50 kDa was detected only in cell lysates, while the supernatant was negative as shown in Fig. 4B (lanes 2, 4). To confirm its cellular localization, stable clones were generated by transfecting CA-MMP expression vector into MDCK cells and immunostained with M2 monoclonal antibody against the FLAG tag. As shown in Fig. 4C, intense signals were detected intracellularly in CA-MMP, but not the control transfected cells. Thus, CA-MMP appears to be mainly a cell-associated product, not readily secreted extracellularly.

3.5. Expression and characterization of CA-MMP active proteinase

To circumvent the difficulty associated with purification of a cell-associated protein, a chimera strategy was employed to fuse the signal peptide and prodomain including the RXKR motif of stromelysin-3 with the catalytic domain of CA-MMP. As illustrated in Fig. 5A, this chimera construct is capable of directing the synthesis of a fusion protein in the ER, which will undergo furin mediated zymogen activation as reported [23]. In addition, a FLAG tag was engineered to the C-terminus to facilitate identification and purification as described [21]. Indeed, when a stable cell line was generated by transfecting the chimera construct into MDCK cells, a 32 kDa protein product can be detected from its supernatant in large quantity by Western blotting using M2 antibody (data not shown). Conditioned media from this stable cell line were subsequently collected and chromatographed on an M2 affinity column (0.5×2 cm). As shown in Fig. 5B, a protein species around 32 kDa in the supernatant (lane 2) was absorbed to the column and could be eluted with FLAG peptide at a 1:1 molar ratio (lane 4). A expected, there is a concomitant decrease in intensity for the protein species around 32 kDa in the flow-through fraction (Fig. 5B, lane 3). The purified CA-MMP can be detected by both M2 antibody or anti-CA-MMP specific antibody on Western blots (Fig. 5B, lane 5, and data not shown). Since almost all MMPs can be assayed for activity by zymography, the same preparation of purified CA-MMP was analyzed by gelatin zymography. As shown in lane 6 of Fig. 5B, a single gelatinolytic species was observed in a molecular weight consistent with the purified material. Thus, it is concluded that the CA-MMP catalytic domain can be expressed and purified in an active form employing this chimera strategy.

CA-MMP catalytic domain expresses gelatinolytic activity sensitive to MMP inhibitors. Since zymography is a technique capable of detecting both latent as well as active MMPs, it is not clear whether the purified CA-MMP is fully active as predicted by the designing strategy. To this end, the proteolytic potential of the purified CA-MMP is characterized against denatured type I collagen (gelatin) in a solution assay. As shown in Fig. 5C, CA-MMP cleaves gelatin efficiently (lane 2). Consistent with its homology to MMPs, the gelatinolytic activity of CA-MMP is inhibited by MMP inhibitors both synthetic (BB94) or natural (TIMP2) (Fig. 5C, lanes 3 and 4). Although the cleaved sites on gelatin have not been identified, CA-MMP appears to cleave it in a novel pattern different from other MMPs such as gelatinase A or B, perhaps reflecting a unique catalytic pocket.

4. Discussion

CA-MMP is apparently the mouse homologue of human MMP21/22 recently reported by Gururajan et al. and MMP23 by Velasco et al. [17–19]. While both papers described the cloning, chromosomal localization as well as tissue distribution of the human gene, its enzymic property remains to be defined. Attempts to characterize a bacterially derived chimera protein between prodomain of MMP19 and human MMP23 led Velasco and colleagues to conclude that human MMP23 has weak proteolytic activity against a synthetic substrate, but absolutely no activity against protein substrate such as gelatin [19]. This conclusion is in sharp contrast to the finding of this report that CA-MMP encodes an active proteinase capable of degrading large protein substrates in an MMP inhibitor sensitive fashion. It is not clear whether this discrepancy is due to a 14% divergence between human and mouse sequences or experimental strategies for expression. Given the inherent problem of misfolding for mamma-

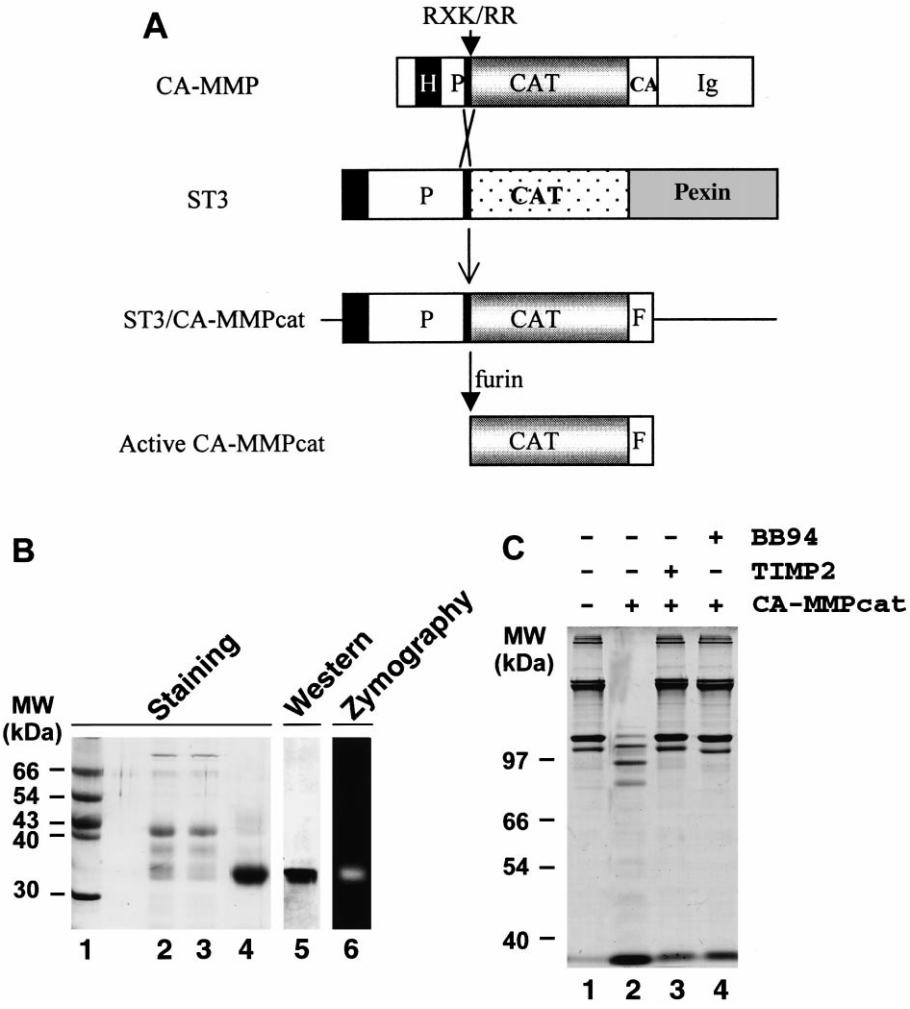


Fig. 5. Expression and characterization of active CA-MMP catalytic domain. A: Expression strategy for CA-MMP catalytic domain. The putative catalytic domain of CA-MMP, Y_{80}-G_{255}, was fused to the prodomain of ST3, M_1-R_{97} to form a chimera ST3/CA-MMPcat as illustrated. A FLAG tag was engineered in-frame to the C-terminus of this construct. This construct is expected to undergo intracellular activation as described for ST3 to generate a fully processed CAM-cat at the furin site. The abbreviations are as in Fig. 4. The block vertical bar represents the RXK/RR motif located between catalytic and prodomains of CA-MMP and ST3. ST3, stromelysin-3/MMP11. B: Purification and characterization of CA-MMPcat. MDCK cell lines stably expressing ST3/CA-MMPcat were generated by transfection, followed by selection in the presence of G418 as described. Collected serum free media (lane 2) was cleared, passed through a 1 ml M2 agarose column (lane 3). Bound materials were eluted with free FLAG peptide competitively and analyzed by staining (lane 4, 500 ng/lane), immunoblotting (lane 5, 25 ng/lane) and zymography (lane 6, 25 ng). C: Proteolytic activity of CA-MMPcat. Denatured type I collagen (5 μg in a 20 μl reaction) was incubated alone (lane 1) or with purified CA-MMPcat (lanes 2–4, 20 ng/lane) at 37°C for 6 h in 50 mM Tris-HCl, 150 mM NaCl, 5 mM $CaCl_2$, in the absence (lanes 1, 2) or presence of metalloproteinase inhibitors BB94 (lane 4, 5 μM) or TIMP2 (lane 3, 50 ng). The reaction mixtures were analyzed by SDS-PAGE as described [22].

lian proteins expressed in *E. coli*, characterization of the human enzyme produced in a mammalian system will help to clarify the differences.

The physiological function of CA-MMP is presently unknown. Expression profile of its human counterpart demonstrated a close association of CA-MMP with reproductive processes based on high levels of expression in tissues such as ovaries, testis and prostates [19]. While the expression of CA-MMP in mouse reproductive tissues has not been determined, it is expressed at relatively high level in heart and lung tissues of mouse, consistent with findings in human tissue [19]. In contrast, mouse spleen expresses significant amount of CA-MMP while human spleen is basically negative. This difference may reflect the detection sensitivity of the methods used, i.e. RT-PCR vs. Northern blotting. While its expression during human embryonic development has not been established, mouse embryo at day 7 expresses CA-MMP at relatively high level when compared to those at stages day 11, 15 and 17, suggesting that CA-MMP may play a critical role during early embryonic development. These data suggest a diverse picture of CA-MMP/MMP21/22/23 expression in both embryonic stages as well as adult tissues, thus, a potentially diverse biological function as well.

CA-MMP is a cell-associated proteinase. Based on immunoprecipitation and immunofluorescence studies (Fig. 4), CA-MMP appears to be associated with cells intracellularly with undetectable secretion. The mechanism of cellular localization is not clear at present. While the hydrophobic region at the N-terminus is consistent with that of a signal peptide, it is possible that it may serve as a transmembrane domain, thus an-

choring CA-MMP as a type II transmembrane proteinase. Alternatively, the novel Ig-fold and cysteine array domains may render CA-MMP membrane- or cell-associated. Further studies are in progress to distinguish these possibilities.

The discovery of CA-MMP demonstrates a divergent and alternative pathway for the evolution of novel members in the MMP superfamily with the following implications. First of all, the lack of a classic cysteine switch and the conserved hemopexin-like domain suggest that CA-MMP branched out of the main MMP tree early in evolution. Indeed, a similar MMP-like gene without a cysteine switch and a hemopexin-like domain has been identified in this laboratory from the *C. elegans* genome (unpublished data). Should CA-MMP or its relatives evolve in parallel to the main MMP branch, it will be interesting to see whether this alternative pathway has been as successful as the classic MMPs with more than 20 members [4]. The recent explosion of new MMP discoveries could be accredited in part to degenerate RT-PCR using primers against the conserved PRCGXPD and HEXGH motifs [10,32–35]. This strategy will apparently fail to identify CA-MMP-like molecule for the absence of a PRCGXPD motif. Thus, new strategies have to be devised to uncover new CA-MMP family members. With the rapid progress of the genome projects on human, mouse as well as other species, a definitive answer will emerge in the near future. Secondly, proprotein convertase mediated MMP activation/processing may be a mechanism of activation evolved as early as MMP catalysis. Out of all motifs and domains found in the classic MMPs, CA-MMP possesses only the catalytic domain and the RXK/RR activation/processing motif. Recognized originally in MMP11 as the signal for zymogen activation [23], a similar RXK/RR motif is present in MMP14–17 (MT1,2,3,4-MMPs) and the newly identified MMP24/MT5-MMP [5,10,20,33–35]. Since the MT-MMPs are known to activate other MMPs without the RXK/RR motifs such as MMP2 and MMP13 [10,36], the proprotein convertase mediated zymogen activation pathway is becoming the dominant mechanism governing MMP activities leading to ECM breakdown. The presence of an RXK/RR motif in a gene such as CA-MMP which branched out early in evolution suggests that the proprotein convertase pathway might have been the method of choice for zymogen activation, even before the cysteine switch latency was evolved. The validity of this argument awaits the identification of more CA-MMP-like genes, which have RXK/RR motifs, but without the cysteine switch. Thirdly, the cysteine array and Ig-like C2 fold domain located at the carboxyl terminus of CA-MMP join the repertoire of MMP substructures such as hemopexin-like domain. In an analogy to the classic MMPs, these two novel domains may specify substrate specificity and inhibitor bindings for CA-MMP. Alternatively, these two domains may determine cellular localization as the full-length protein is cell-associated (Fig. 4). Efforts are underway to uncover the contribution of these two motifs to CA-MMP function biochemically. Finally, CA-MMP may possess a novel mechanism of latency. The presence of an RXK/RR motif suggests that the prodomain of CA-MMP will be cleaved off by furin or related enzymes in the trans-Golgi network. Given the fact that all MMPs discovered so far are encoded as zymogens, it is likely that CA-MMP is synthesized as a zymogen and processed into active form by furin or furin-like enzymes [23,37]. Since CA-MMP has no cysteine residue in the prodomain, it may have a novel mechanism of latency. Elucidation of its latency mechanism will broaden our understanding of structural/functional relationships for the MMP gene family. Taken together, CA-MMP along with its human counterpart MMP21/22/23 helps define a new subfamily of MMPs with distinct structures and potential functional consequences in ECM remodeling.

Acknowledgements: This study was supported in part by grant CA76308 from the National Cancer Institute, American Heart Association Grant-in-Aid 9750197N, Elsa Pardee Foundation, University of Minnesota grant-in-aid, Minnesota Medical Foundation and Arthritis Foundation (MN chapter). The author wishes to thank Jun Yi for technical help; R. Fridman for TIMP2; Stephen J. Weiss, Hideaki Nagase and Fred Woessner for thoughtful inputs and comments on this study.

References

[1] Werb, Z. and Chin, J.R. (1998) Ann. N.Y. Acad. Sci. 857, 110–118.
[2] Chambers, A.F. and Matrisian, L.M. (1997) J. Natl. Cancer Inst. 89, 1260–1270.
[3] Crawford, H.C. and Matrisian, L.M. (1994) Invasion Metastasis 14, 234–245.
[4] Massova, I., Kotra, L.P., Fridman, R. and Mobashery, S. (1998) FASEB J. 12, 1075–1095.
[5] Nagase, H. (1997) Biol. Chem. 378, 151–160.
[6] Birkedal-Hansen, H., Moore, W.G., Bodden, M.K., Windsor, L.J., Birkedal-Hansen, B., DeCarlo, A. and Engler, J.A. (1993) Crit. Rev. Oral Biol. Med. 4, 197–250.
[7] Weiss, S.J. (1989) N. Engl. J. Med. 320, 365–376.
[8] Matrisian, L.M., Bowden, G.T., Krieg, P., Furstenberger, G., Briand, J.P., Leroy, P. and Breathnach, R. (1986) Proc. Natl. Acad. Sci. USA 83, 9413–9417.
[9] Huhtala, P., Chow, L., Shows, T. and Tryggvason, K. (1992) Matrix 1 (Suppl.), 84.
[10] Sato, H., Takino, T., Okada, Y., Cao, J., Shinagawa, A., Yamamoto, E. and Seiki, M. (1994) Nature 370, 61–65.
[11] Apte, S.S., Fukai, N., Beier, D.R. and Olsen, B.R. (1997) J. Biol. Chem. 272, 25511–25517.
[12] Van Wart, H.E. and Birkedal-Hansen, H. (1990) Proc. Natl. Acad. Sci. USA 87, 5578–5582.
[13] Weiss, S.J., Peppin, G., Ortiz, X., Ragsdale, C. and Test, S.T. (1985) Science 227, 747–749.
[14] Desrochers, P.E., Mookhtiar, K., Van Wart, H.E., Hasty, K.A. and Weiss, S.J. (1992) J. Biol. Chem. 267, 5005–5012.
[15] Nagase, H., Enghild, J.J., Suzuki, K. and Salvesen, G. (1990) Biochemistry 29, 5783–5789.
[16] Sang, Q.X., Birkedal-Hansen, H. and Van Wart, H.E. (1995) Biochim. Biophys. Acta 1251, 99–108.
[17] Gururajan, R., Lahti, J.M., Grenet, J., Easton, J., Gruber, I., Ambros, P.F. and Kidd, V.J. (1998) Genome Res. 8, 929–939.
[18] Gururajan, R., Grenet, J., Lahti, J.M. and Kidd, V.J. (1998) Genomics 52, 101–106.
[19] Velasco, G., Pendas, A.M., Fueyo, A., Knauper, V., Murphy, G. and Lopez-Otin, C. (1999) J. Biol. Chem. 274, 4570–4576.
[20] Pei, D. (1999) J. Biol. Chem. 274, 8925–8932.
[21] Pei, D. and Yi, J. (1998) Protein Expr. Purif. 13, 277–281.
[22] Pei, D., Majmudar, G. and Weiss, S.J. (1994) J. Biol. Chem. 269, 25849–25855.
[23] Pei, D. and Weiss, S.J. (1995) Nature 375, 244–247.
[24] Harlow, E. and Lane, D. (1988) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[25] Santavicca, M. et al. (1996) Biochem. J. 315, 953–958.
[26] Pei, D. and Weiss, S.J. (1996) J. Biol. Chem. 271, 9135–9140.
[27] Stocker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Ruth, F.X., McKay, D.B. and Bode, W. (1995) Protein Sci. 4, 823–840.
[28] Bajorath, J., Peach, R.J. and Linsley, P.S. (1994) Protein Sci. 3, 2148–2150.
[29] Cho, K.R. et al. (1994) Genomics 19, 525–531.

[30] Vigers, G.P., Anderson, L.J., Caffes, P. and Brandhuber, B.J. (1997) Nature 386, 190–194.
[31] Korinek, V., Stefanova, I., Angelisova, P., Hilgert, I. and Horejsi, V. (1991) Immunogenetics 33, 108–112.
[32] Yang, M., Hayashi, K., Hayashi, M., Fujii, J.T. and Kurkinen, M. (1996) J. Biol. Chem. 271, 25548–25554.
[33] Takino, T., Sato, H., Shinagawa, A. and Seiki, M. (1995) J. Biol. Chem. 270, 23013–23020.
[34] Puente, X.S., Pendas, A.M., Llano, E., Velasco, G. and Lopez-Otin, C. (1996) Cancer Res. 56, 944–949.
[35] Will, H. and Hinzmann, B. (1995) Eur. J. Biochem. 231, 602–608.
[36] Knauper, V., Will, H., Lopez-Otin, C., Smith, B., Atkinson, S.J., Stanton, H., Hembry, R.M. and Murphy, G. (1996) J. Biol. Chem. 271, 17124–17131.
[37] Steiner, D.F. (1998) Curr. Opin. Chem. Biol. 2, 31–39.