Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 2800–2804, April 1990
Biochemistry

cDNA cloning and expression of a metalloproteinase inhibitor related to tissue inhibitor of metalloproteinases

(extracellular matrix/gene family/collagenase)

THOMAS C. BOONE*, MERRIE JO JOHNSON*, YVES A. DE CLERCK†, AND KEITH E. LANGLEY*

*Amgen, Inc., Thousand Oaks, CA 91320; and †Division of Hematology/Oncology, Childrens Hospital of Los Angeles and the University of Southern California, Los Angeles, CA 90027

Communicated by Fred Sherman, January 16, 1990 (received for review September 28, 1989)

ABSTRACT **The purification and characterization of a metalloproteinase inhibitor (MI) from bovine aortic endothelial cells, and the demonstration that it is related to, but distinct from, tissue inhibitor of metalloproteinases (TIMP), have previously been reported [De Clerck, Y. A., Yean, T.-D., Ratzkin, B. J., Lu, H. S. & Langley, K. E. (1989) *J. Biol. Chem.* 264, 17445–17453]. The cDNA cloning of the bovine MI and its human homolog is now reported. The bovine cDNA cloning used probes designed on the basis of NH_2-terminal amino acid sequence of bovine MI. The human cDNA cloning in turn used probes representing parts of the bovine cDNA nucleotide sequence. Both cDNAs encode leader sequences of 26 amino acids and mature protein sequences of 194 amino acids. The amino acid sequences of the mature proteins are 94% identical. The human MI cDNA was expressed in *Escherichia coli*, and a preparation containing anticollagenase activity was recovered. The amino acid sequence of mature human MI is 38% identical to the sequence for human TIMP, and the 12 cysteines in MI and TIMP are aligned almost identically. Thus MI and TIMP comprise an inhibitor family.**

The extracellular matrices are complex structures that contain collagens, proteoglycans, glycoproteins (fibronectin, laminin), and in some cases, elastin (see review, ref. 1). Several related matrix metalloproteinases (MMPs), members of the collagenase gene family, have been implicated in the degradation and turnover of these constituents (see review, ref. 2). These zinc-dependent endopeptidases include interstitial collagenase (MMP-1), 72-kDa gelatinase (type IV collagenase, MMP-2), 92-kDa gelatinase, and stromelysin (proteoglycanase, transin, MMP-3). Virtually all extracellular matrix constituents can be degraded by this family of enzymes, and while there is some specificity in substrate recognition, there is also considerable overlap (2).

The MMPs are synthesized and secreted as inactive zymogens, and their activity in the extracellular milieu is regulated by activators and inhibitors (2). Changes in the balance between active enzymes and their inhibitors may be involved in a number of human diseases, including arthritis, fibrosis, and tumor invasion (see review, ref. 3).

Several inhibitors of metalloproteinases have been described and characterized. These include small cationic proteins extracted from cartilage and large blood vessels (4); α_2-macroglobulin, a major serum inhibitor with $M_r \approx 725{,}000$ (5); and tissue inhibitor of metalloproteinases (TIMP), a glycoprotein of $M_r \approx 30{,}000$ (6, 7) for which the human cDNA has been cloned and sequenced (8, 9). TIMP is ubiquitous among tissues and species and is considered a major regulatory inhibitor in the extracellular milieu (see review, ref. 2).

We recently described the purification and characterization of TIMP and a novel metalloproteinase inhibitor (designated MI), both secreted by bovine aortic endothelial cells (10). The MI was shown to have an apparent M_r of 27,500, to be unglycosylated, and to have 51% identity to the bovine TIMP in the first 45 amino acids. The findings indicated that TIMP and MI are related but distinct and that a family of metalloproteinase inhibitors exists. The findings also raised the possibility that there could be a homolog of MI in the human species. In this paper, we report the cloning and sequencing of cDNAs for bovine MI and human MI and expression of the human MI cDNA in *Escherichia coli*.‡

MATERIALS AND METHODS

Isolation and Characterization of a Bovine cDNA Clone. A λgt11 cDNA library made using mRNA isolated from bovine aortic endothelial cells was purchased from Clontech. Approximately 10^6 phage were plated, lifted onto GeneScreen-*Plus* (DuPont/NEN) membranes in duplicate, and hybridized independently to each of two degenerate ^{32}P-phosphorylated oligodeoxynucleotide probe mixtures. These probe mixtures were designed for recognition of the sense strand of DNA corresponding to the NH_2-terminal amino acids 21–30 and 32–41 of bovine MI (10) (underlined in Fig. 2) and were as follows: probe 1, 5′-GTC-IAC-YTC-YTT-YTT-GTT-IAC-IGC-YTT-IGC-3′; probe 2, 5′-CTT-IAT-IGG-RTT-ICC-RTA-IAT-RTC-RTT-ICC-3′. The symbols Y (= C or T) and R (= A or G) indicate positions where degeneracy was introduced into the mixtures; i.e., each mixture contained 16 oligomers of 30 nucleotides in length. The mixtures represented regions where the codons of the corresponding amino acids have relatively little degeneracy and where there was the least possible homology to the reported nucleotide sequence for TIMP cDNA (8, 9). Deoxyinosine (I) was used at codon nucleotide positions of fourfold degeneracy (11) in order to limit the number of oligonucleotides in each probe mixture. Hybridizations were done overnight at 50–55°C in 6× SSC (SSC is 15 mM sodium citrate/150 mM sodium chloride, pH 7) containing 5× Denhardt's solution (12), 0.5% (wt/vol) SDS, and 50 μg of sheared and denatured herring sperm DNA per ml. The filters were washed in 6× SSC/0.5% SDS at ≈55°C. One clone was obtained that hybridized to both probes. Nucleotide sequencing was performed by cloning a restriction fragment of the bovine MI clone from λgt11 into pUC19 and then subcloning into M13 vectors followed by primed DNA synthesis on single-stranded DNA templates in the presence of dideoxynucleoside triphosphates (13) and Sequenase (United States Biochemical).

Isolation and Characterization of Human cDNA Clones. A λgt11 cDNA library made using mRNA isolated from human

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: MI, metalloproteinase inhibitor; MMP, matrix metalloproteinase; TIMP, tissue inhibitor of metalloproteinases.

‡The sequences reported in this paper have been deposited in the GenBank data base (accession no. M32303 for the bovine sequence and M32304 for the human sequence).

heart tissue (fetal aorta) was also purchased from Clontech and screened as described above, with two separate pairs of ^{32}P-phosphorylated oligonucleotide probes of 51 nucleotides in length. The probes were synthesized as exact matches to regions of the antisense strand of the bovine MI cDNA. One pair contained probes 1 and 2, and the second pair contained probes 3 and 4, as identified in the legend to Fig. 2. These cDNA regions have relatively low homology to the corresponding regions of TIMP cDNA (8, 9). Three clones were obtained that hybridized to both probe pairs, and one clone was sequenced as described above.

Expression of Recombinant Human MI in *E. coli*. The mature human MI protein was expressed in *E. coli* by utilizing an *Nco* I site at the codon for amino acid 1 of the leader sequence, a *Bam*HI site at the codon for amino acid 42 of the mature protein, and a *Stu* I site three nucleotides downstream from the termination codon (see Figs. 1 and 2). The fragment from *Nco* I to *Stu* I was first cloned in an expression plasmid, pCFM1156pL, from *Nco* I to *Sst* II (the *Sst* II site had been blunted using the Klenow fragment of DNA polymerase I) to generate p1156HMI1 (Fig. 1). pCFM1156pL contains the P_L promoter/operator region of phage λ and has a temperature-sensitive replicon; it is a derivative of pCFM4722 (14) having the naturally occurring λpL promoter sequence instead of the synthetically altered sequence present in pCFM4722. It was necessary to eliminate leader sequences from the cDNA in p1156HMI1 and to introduce sequences appropriate for expression of the mature protein in *E. coli*. Therefore a synthetic DNA fragment containing a ribosome binding site, an initiation methionine codon, and codons for the first 42 amino acids of the mature human MI was cloned from *Xba* I to *Bam*HI, first into M13 mp11 to confirm the sequence and then into p1156HMI1 to generate p1156HMI2 (Fig. 1).

Plasmid p1156HMI2 was then put into an *E. coli* host strain, FM5, that harbors the gene for the λ cI857 temperature-sensitive repressor on the chromosome. Increasing the temperature from 30°C to 42°C during exponential growth induces expression of the human MI gene.

Recovery of *E. coli*-Expressed Recombinant Human MI. Strain FM5 transformed with the p1156HMI2 plasmid was grown in shaker flasks in medium containing 16 g of Bactotryptone, 10 g of yeast extract, and 5 g of NaCl per liter. Temperature induction was carried out at OD_{600} of 0.6, and the OD_{600} at harvesting was 2.7. Approximately 12 g (wet weight) of harvested cells was suspended in water and broken by two passes through a French pressure cell. The cell lysate was centrifuged at 15,000 × *g* for 30 min, and the pellet fraction was suspended in 10 ml of 50 mM Tris·HCl/50 mM dithiothreitol/2% (wt/vol) sodium *N*-lauroylsarcosine, pH 8.5. After incubation at 50°C for 10 min, and room temperature for 2 hr, the mixture was clarified by centrifugation at 20,000 × *g* for 30 min. The supernatant was applied to a Sephacryl S-200 gel filtration column (Pharmacia; 2.5 × 54 cm) equilibrated in 20 mM Tris·HCl/1% sodium *N*-lauroylsarcosine, pH 8, at room temperature. Fractions of 2.9 ml were collected at a flow rate of 14 ml/hr and analyzed by A_{280} and by SDS/polyacrylamide gel electrophoresis (PAGE). Pools were made of fractions 42–50 (void-volume region; no detectable MI) and of fractions 65–75 (included region; containing MI). These pools were dialyzed over a 3-day period against 20 mM Tris·HCl (pH 8) at 4°C. The dialyzed samples were concentrated to about 5 ml in an Amicon stirred cell with a YM10 membrane and filtered through 0.45-μm filters.

Collagenase Inhibition Assays. Assays for inhibition of the degradation of type I rat skin collagen by rabbit synovial fibroblast collagenase were performed according to the film assay described by Johnson-Wint (15). Sources of collagen

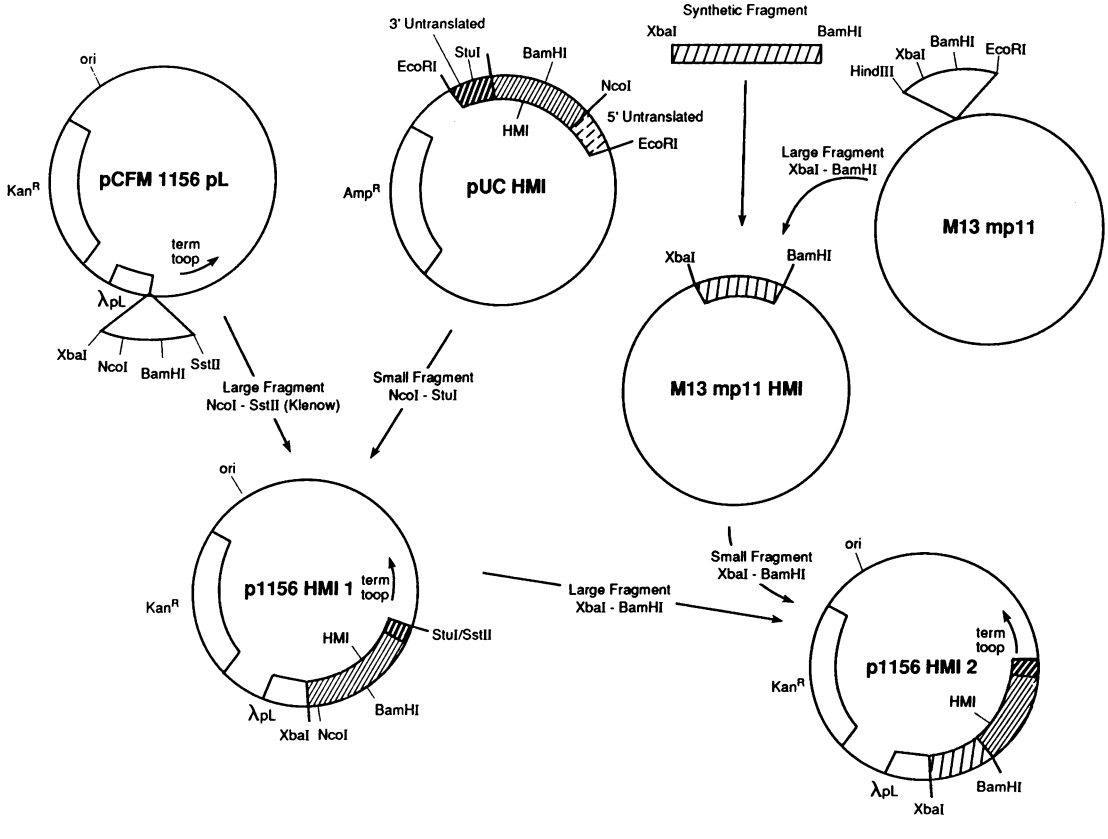

FIG. 1. Construction of plasmid p1156HMI2 for expression of human MI in *E. coli*, as described in *Materials and Methods*. Term toop refers to the transcription termination signal from λ, and HMI refers to human MI cDNA.

and collagenase, other details of the assay procedure, and calculation of units have been presented (10).

Polyacrylamide Gel Electrophoresis. SDS/PAGE was carried out by the method of Laemmli (16). Stacking gels contained 4% (wt/vol) acrylamide, and separating gels contained 12.5% acrylamide.

SDS/gelatin/PAGE was used to test the antigelatinolytic activity of protein bands. This method has been described in detail (10, 17). In brief, gelatin is incorporated into the gels, and after electrophoresis of samples, the gels are incubated in the presence of collagenase/gelatinase to digest gelatin. The gels are then stained with Coomassie blue and bands with antigelatinolytic activity are evident as blue zones representing undigested gelatin.

RESULTS

Isolation and Sequencing of a Bovine MI cDNA Clone. Since the bovine MI had been purified from medium conditioned by bovine aortic endothelial cell line $NCACl_2$ (10), a cDNA library made from bovine aortic endothelial cell mRNA was probed for the cDNA (*Materials and Methods*). A single clone that hybridized to both probes was obtained, and its sequence is shown in Fig. 2. The sequence codes for a mature

```
1 Bovine MI cDNA   GA ATT CCG GCT TCT ATG GAG CAC TCG GGA CCA GGT CCG CGG CGC GCG CAC TCG CTC GCT CGC CGC CCC CCA GCC AGC TCT CGC TTC CGC GCC GCC AGC CGC GCC CCG CGC  110
2 Human MI cDNA                                            GA ATT CCG GCC CGC CGT CCC CCA CCC CGC CGC CCC GCC CGG CGA ATT GCG CCC CGC GCC CCT CCC CTC GCG CCC   74

1 CTC CTC GCT GCA CCC CGC GAC CTA GAG CCA AGA AAG TTT GTG TGG CGA GTG AGG GCC GGA GAG GAG AGC GCG CCC GCG GAG TGC CGT CCA GAC CAG CGC GGC CCC GGC GGA GAG GGG AGC GCC  233
2 CCG AGA CAA AGA GGA GAG AAA GTT TGC GCG GCC GAG CGG GGC AGG TGA GGA GGG TGA GCC GCG CGG GAG GGG CCC GCC TCG GCC CCG GCT CAG CCC CCG CCC GCG CCC CCA GCC CGC CGC CGC  197

1 CCG AGC CCA GGC GGC GGC GGC TAG CCC GAG TCC GCG ACC CCC GCC CCT CCG CCC GCC ATG GGC GCC GCC GCC CGC AGC CTG CCG CTC GCG TTC TGC CTC CTG CTG CTG GGG ACG CTG CTC CCC  356
2 GAG CAG CGC CCG GAC CCC CCA GCG GCG GCC CCC GCC CGC CCA GCC CCC CGG CCC GCC ATG GGC GCC GCG GCC CGC ACC CTG CGG CTG GCG CTC GGC CTC CTG CTG CTG GCG ACG CTG CTT CGC  320
                                                                                            -20                         -10
3 Bovine MI protein                                          Met Gly Ala Ala Ala Arg Ser Leu Pro Leu Ala Phe Cys Leu Leu Leu Leu Gly Thr Leu Leu Pro
4 Human MI protein                                           Met Gly Ala Ala Ala Arg Thr Leu Arg Leu Ala Leu Gly Leu Leu Leu Leu Ala Thr Leu Leu Arg
5 Human TIMP protein                                         Met  -  Ala Pro Phe Glu Pro Leu  -  Ala Ser Gly Ile Leu Leu Leu Leu  -  Trp Leu Ile Ala

1 CGG GCC GAC GCC TGC AGC TGC TCC CCG GTG CAC CCG CAA CAG GCG TTT TGC AAT GCA GAC ATA GTG ATC AGG GCC AAA GCA GTC AAT AAG AAG GAG GTG GAC TCT GGC AAC GAC ATC TAC GGC  479
2 CCG GCC GAC GCC TGC AGC TGC TCC CCG GTG CAC CCG CAA CAG GCG TTT TGC AAT GCA GAT GTA GTG ATC AGG GCC AAA GCG GTC AGT GAG AAG GAA GTG GAC TCT GGA AAC GAC ATT TAT GGC  443
                 -1  1                       10                                  20                                  30
3 Arg Ala Asp Ala [Cys] Ser [Cys] Ser Pro Val His Pro Gln Gln Ala Phe [Cys] Asn Ala Asp Ile Val Ile Arg Ala Lys Ala Val Asn Lys Lys Glu Val Asp Ser Gly Asn Asp Ile Tyr Gly
4 Pro Ala Asp Ala [Cys] Ser [Cys] Ser Pro Val His Pro Gln Gln Ala Phe [Cys] Asn Ala Asp Val Val Ile Arg Ala Lys Ala Val Ser Glu Lys Glu Val Asp Ser Gly Asn Asp Ile Tyr Gly
5 Pro Ser Arg Ala [Cys] Thr [Cys] Val Pro Pro His Pro Gln Thr Ala Phe [Cys] Asn Ser Asp Leu Val Ile Arg Ala Lys Phe Val Gly Thr Pro Glu Val Asn Gln Thr Thr Leu Tyr Gln Arg   37

1 AAC CCC ATC AAG CGG ATT CAG TAT GAG ATC AAG CAG ATA AAG ATG TTC AAG GGA CCT GAT CAG GAC ATA GAG TTT ATC TAC ACA GCC CCC GCC GCT GCC GTG TGT GGG GTC TCG CTG GAC ATT  602
2 AAC CCT ATC AAG AGG ATC CAG TAT GAG ATC AAG CAG ATA AAG ATG TTC AAA GGG CCT GAG AAG GAT ATA GAG TTT ATC TAC ACG GCC CCC TCC TCG GCA GTG TGT GGG GTC TCG CTG GAC GTT  566
      40                            50                                  60                              70
3 Asn Pro Ile Lys Arg Ile Gln Tyr Glu Ile Lys Gln Ile Lys Met Phe Lys Gly Pro Asp Gln Asp Ile Glu Phe Ile Tyr Thr Ala Pro Ala Ala Ala Val [Cys] Gly Val Ser Leu Asp Ile
4 Asn Pro Ile Lys Arg Ile Gln Tyr Glu Ile Lys Gln Ile Lys Met Phe Lys Gly Pro Glu Lys Asp Ile Glu Phe Ile Tyr Thr Ala Pro Ser Ser Ala Val [Cys] Gly Val Ser Leu Asp Val
5 Tyr Glu Ile Lys  -   -  Met Thr Lys Met Tyr Lys Gly Phe Gln Ala Leu Gly Asp Ala Ala Asp Ile Arg Phe Val Tyr Thr Pro Ala Met Glu Ser Val [Cys] GlyTyrPheHisArg Ser His   77

1 GGA GGA AAG AAG GAG TAT CTC ATT GCA GGG AAG GCC GAG GGG AAT GGC AAT ATG CAT ATC ACC CTC TGT GAC TTC ATC GTG CCC TGG GAC ACC CTG AGT GCC ACC CAG AAG AAG AGC CTG AAC  725
2 GGA GGA AAG AAG GAA TAT CTC ATT GCA GGA AAG GCC GAG GGG GAC GGC AAG ATG CAC ATC ACC CTC TGT GAC TTC ATC GTG CCC TGG GAC ACC CTG AGC ACC ACC CAG AAG AAG AGC CTG AAC  689
      80                            90                                  100                              110
3 Gly Gly Lys Lys Glu Tyr Leu Ile Ala Gly Lys Ala Glu Gly Asn Gly Asn Met His Ile Thr Leu [Cys] Asp Phe Ile Val Pro Trp Asp Thr Leu Ser Ala Thr Gln Lys Lys Ser Leu Asn
4 Gly Gly Lys Lys Glu Tyr Leu Ile Ala Gly Lys Ala Glu Gly Asp Gly Lys Met His Ile Thr Leu [Cys] Asp Phe Ile Val Pro Trp Asp Thr Leu Ser Thr Thr Gln Lys Lys Ser Leu Asn
5 Asn Arg Ser Glu Glu Phe Leu Ile Ala Gly Lys Leu Gln  -  Asp Gly Leu Leu His Ile Thr Thr [Cys] Ser Phe Val Ala Pro Trp Asn Ser Leu Ser Leu Ala Gln Arg Arg Gly Phe Thr  117

1 CAC AGG TAC CAG ATG GGC TGT GAG TGC AAG ATC ACT CGA TGC CCC ATG ATC CCA TGC TAC ATC TCC TCT CCG GAC GAG TGC CTC TGG ATG GAC TGG GTC ACG GAG AAG AAC ATC AAC GGA CAC  848
2 CAC AGG TAC CAG ATG GGC TGC GAG TGC AAG ATC ACG CGC TGC CCC ATG ATC CCG TGC TAC ATC TCC TCC CCG GAC GAG TGC CTC TGG ATG GAC TGG GTC ACA GAG AAG AAC ATC AAC GGG CAC  812
  120                           130                                 140                                 150                                   160
3 His Arg Tyr Gln Met Gly [Cys] Glu [Cys] Lys Ile Thr Arg [Cys] Pro Met Ile Pro [Cys] Tyr Ile Ser Ser Pro Asp Glu [Cys] Leu Trp Met Asp Trp Val Thr Glu Lys Asn Ile Asn Gly His
4 His Arg Tyr Gln Met Gly [Cys] Glu [Cys] Lys Ile Thr Arg [Cys] Pro Met Ile Pro [Cys] Tyr Ile Ser Ser Pro Asp Glu [Cys] Leu Trp Met Asp Trp Val Thr Glu Lys Asn Ile Asn Gly His
5 Lys Thr Tyr Thr Val Gly [Cys]GluGlu[Cys]Thr Val Phe Pro [Cys] Leu Ser Ile Pro [Cys] Lys Leu Gln Ser Gly Thr His [Cys] Leu Trp Thr Asp Gln Leu Leu Gln Gly Ser Glu Lys Gly Phe  159

1 CAG GCC AAG TTC TTC GCC TGC ATC AAG AGA AGC GAC GGC TCC TGC GCC TGG TAC CGC GGA GCA GCA CCC CCC AAG CAG GAG TTT CTG GAC ATC GAG GAC CCG TAA GCA GGC CAC CAG GAC TCC  971
2 CAG GCC AAG TTC TTC GCC TGC ATC AAG AGA AGT GAC GGC TCC TGT GCG TGG TAC CGC GGC GCG GCG CCC CCC AAG CAG GAG TTT CTC GAC ATC GAG GAC CCA TAA GCA GGC CTC CAA CGC CCC  935
                                        170                                 180                             190
3 Gln Ala Lys Phe Phe Ala [Cys] Ile Lys Arg Ser Asp Gly Ser [Cys] Ala Trp Tyr Arg Gly Ala Ala Pro Pro Lys Gln Glu Phe Leu Asp Ile Glu Asp Pro
4 Gln Ala Lys Phe Phe Ala [Cys] Ile Lys Arg Ser Asp Gly Ser [Cys] Ala Trp Tyr Arg Gly Ala Ala Pro Pro Lys Gln Glu Phe Leu Asp Ile Glu Asp Pro
5 Gln Ser Arg His Leu Ala [Cys] Leu Pro Arg Glu Pro Gly Leu [Cys] Thr Trp Gln Ser Leu Arg Ser Gln Ile Ala                                                        184

1 TGG GGC CAA TTG ACA GTG TCC AAG AGT TCA GAC TGG TCC AGC TCC GAC ATC CCT TCC TGG ACA CAG CAT GAA TAA A                                              1047
2 TGT GGC CAA CTG CAA AAA AAG CCT CCA AGG GTT TCG ACT GGT CCA GCT CTG ACA TCC CTT CCT GGA AAC AGC ATG AAT AAA ACA CTC ATC CCC GGA ATT C           1035
```

FIG. 2. Nucleotide sequences of cloned cDNAs, and deduced amino acid sequences. Lines 1 and 2 show the nucleotide sequences for the bovine MI cDNA and human MI cDNA, as indicated. Lines 3 and 4 show the amino acid sequences for bovine MI and human MI deduced from the corresponding cDNA sequences. Line 5 shows the amino acid sequence of human TIMP, deduced from cDNA (8, 9). The numbers above line 3 refer to the amino acids in bovine MI (line 3) and human MI (line 4); the designation of the indicated cysteine as residue 1 is based on the known NH_2-terminal sequences of mature bovine MI (10) and mature human TIMP (7, 8); the negative numbers refer to the deduced leader sequences. Amino acids in the human TIMP sequence (line 5) have been slightly realigned in places relative to the MI amino acid sequences in order to maximize identity and retain precise alignment of cysteine residues; dashes in the human TIMP sequence do not indicate missing residues, but rather places where realignment has been done. Cysteine residues are enclosed in boxes to highlight their alignment in the three sequences. Numbers to the right of lines 1 and 2 indicate nucleotide numbers, and numbers to the right of line 5 indicate amino acid numbers. The underlined amino acids in line 3 (amino acids 21–30 and 32–41) indicate the sequences from which degenerate oligonucleotide probes were designed for screening the bovine aortic endothelial cell cDNA library (*Materials and Methods*). The underlined nucleotides in line 1 indicate the sequences corresponding to probes used to screen the human fetal aorta cDNA library (*Materials and Methods*; probe 1, nucleotides 435–485; probe 2, nucleotides 702–752; probe 3, nucleotides 810–860; probe 4, nucleotides 900–950). Potential sites of N-linked glycosylation in human TIMP are underlined (line 5, amino acids 30–32 and 78–80), and the poly(A) addition signals (18) in the bovine MI cDNA (line 1, nucleotides 1042–1047) and human MI cDNA (line 2, nucleotides 1011–1016) 3′ untranslated regions are underlined.

protein of 194 amino acids (calculated M_r of 21,718), the first 45 of which exactly match those obtained by NH_2-terminal sequencing of the purified protein (10). A hydrophobic leader sequence (19) of 26 amino acids is also encoded and there is a very G+C-rich 5′ untranslated region. It appears that the 3′ untranslated region was truncated in the cDNA cloning, since there is no *Eco*RI site (used to insert the cDNAs into λgt11) and no sequence after the poly(A) addition signal.

Isolation and Sequencing of a Human MI cDNA Clone. A cDNA library made using mRNA from fetal human heart tissue was probed for a human equivalent of the bovine MI cDNA (*Materials and Methods*). The sequence of one of the isolated clones is shown in Fig. 2. This sequence, like its bovine counterpart, codes for a mature protein of 194 amino acids (calculated M_r of 21,755) and a leader sequence of 26 amino acids. One hundred eighty-three of the 194 amino acids in the mature bovine and human proteins are identical (94% identity), and 20 of the 26 amino acids in the bovine and human leader sequences are identical. As with the bovine cDNA sequence, the 5′ untranslated region in the human cDNA sequence is G+C-rich.

Expression of the Human MI cDNA in *E. coli*. *E. coli* strain FM5/p1156HMI2 was grown with temperature induction and the cells were harvested and broken. After centrifugation of the lysate, the expressed human MI was present in the pellet fraction, in insoluble form. The MI was brought into solution by treatment with *N*-lauroylsarcosine and dithiothreitol and was partially purified by gel filtration in the presence of *N*-lauroylsarcosine. SDS/PAGE of the broken-cell pellet fraction after solubilization with *N*-lauroylsarcosine (Fig. 3*A*, lane 1) and of the MI-containing gel filtration pool after dialysis and concentration (lane 2) gave a major stained band at M_r 24,000–27,000 (relative to the molecular weight markers used) that comigrated with bovine endothelial cell MI (lane 3).

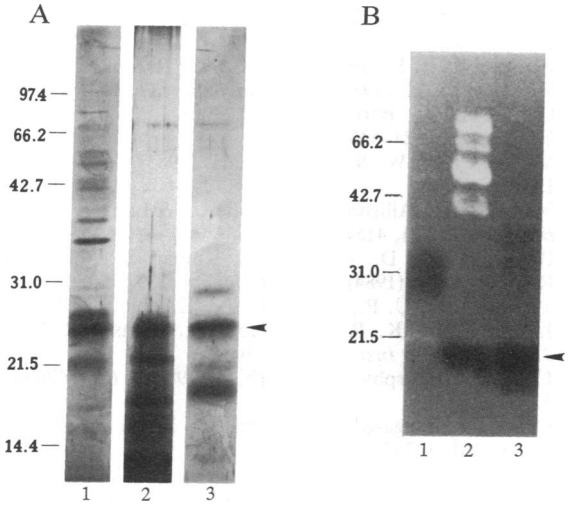

FIG. 3. SDS/PAGE of MI-containing samples. (*A*) Silver staining (20). Lane 1, aliquot (0.05 μl) of sample applied to Sephacryl S-200 gel filtration column after *N*-lauroylsarcosine solubilization of broken-cell pellet fraction from *E. coli* strain FM5/p1156HMI2 expressing human MI; lane 2, aliquot (0.5 μl) of MI-containing pool (dialyzed and concentrated) from the gel filtration column; lane 3, partially purified bovine MI (ref. 10; 0.2 unit). All samples were reduced. Arrowhead indicates the migration position of bovine MI. Positions of molecular weight standards are indicated at left (phosphorylase *b*, M_r 97,400; bovine serum albumin, M_r 66,200; ovalbumin, M_r 42,700; carbonic anhydrase, M_r 31,000; soybean trypsin inhibitor, M_r 21,500; and lysozyme, M_r 14,400). (*B*) SDS/gelatin/PAGE. Lane 1, bovine TIMP (ref. 10; 0.024 unit); lane 2, partially purified bovine MI (ref. 10; 0.05 unit); lane 3, *E. coli*-derived recombinant human MI (≈0.2 μg). The dark zones are indicative of proteins having antigelatinolytic activity. All samples were unreduced since reduction destroys the antigelatinolytic activity of MI and TIMP in this procedure; it should be noted that unreduced MI has a substantially greater mobility than reduced MI in SDS gels (10). Arrowhead indicates the position of bovine MI. The clear zones in lane 2 indicate gelatinolytic proteins present in the partially purified bovine MI. Numbers at left indicate molecular weight markers as for *a*.

The gel filtration void-volume pool, lacking MI, and the included pool containing MI, after dialysis and concentration, were assayed for collagenase inhibition. The void-volume pool, serving as a control for nonspecific effects in the assay, showed no inhibition (aliquots as large as 1 μl were assayed). The MI-containing pool had collagenase inhibition activity of 77.5 units/ml (aliquots of ≥0.5 μl gave complete inhibition). Total protein in the MI-containing pool was about 10 mg, of which roughly 25% was MI, as judged by SDS/PAGE (Fig. 3*A*, lane 2). The MI in the broken cell pellet fraction, after solubilization with *N*-lauroylsarcosine, almost certainly required folding and oxidation (disulfide bond formation) to achieve native, active conformation. After the gel filtration and dialysis procedures described, it is likely that only a portion of the total MI polypeptide had achieved the active conformation. However, this result is sufficient to demonstrate that the expressed polypeptide is a collagenase inhibitor.

We analyzed samples by SDS/gelatin/PAGE, to verify that the *E. coli*-expressed polypeptide responsible for the inhibitory activity had the expected electrophoretic mobility. By this procedure, bands having inhibitory activity toward gelatinolysis can be visualized. *E. coli*-derived human MI samples showed an inhibitory band that comigrated with that for bovine endothelial cell MI (Fig. 3*B*).

To rule out the possibility that the observed activities could be due to *E. coli*-specific material, an included-volume gel filtration pool was obtained from an *E. coli* FM5 strain lacking the p1156HMI2 plasmid. The same methods described for the case of the FM5/p1156HMI2 strain were used. This control sample lacked collagenase inhibition activity and had no bands with antigelatinolytic activity in the SDS/gelatin/PAGE procedure (data not shown).

DISCUSSION

The NH_2-terminal amino acid sequence information obtained for bovine endothelial cell MI (10) was used to obtain a bovine cDNA clone. The cDNA nucleotide sequence codes for a mature protein of 194 amino acids, the first 45 of which exactly match those determined for the endothelial cell protein. Thus the cDNA must certainly represent the bovine MI gene. In turn, we used the bovine cDNA sequence to obtain a human cDNA clone. We conclude that the human cDNA codes for the homolog of bovine MI, based on the high degree of identity (94%) between the deduced amino acid sequences for the mature bovine and human proteins and based on the demonstration of active polypeptide having the expected mobility in SDS/polyacrylamide gels after expression of the human cDNA in *E. coli*.

As in the case of the bovine species (10), it is clear that the human MI is related to, but distinct from, human TIMP. The amino acid sequence of human TIMP, deduced from cDNA nucleotide sequence (8, 9), is shown in Fig. 2 for comparison to the deduced MI amino acid sequences. In each case there is a predicted hydrophobic leader sequence, consistent with these inhibitors being secreted proteins (6, 7, 10). The degree of identity in the sequences for the human MI and human TIMP mature proteins is 38% overall, and higher within certain regions (71% for amino acids 1–24 and 59% for amino acids 83–111). Although the mature human TIMP has 10 fewer amino acids than MI, both contain 12 cysteine residues, an unusually large number, at virtually identical positions. These comparisons clearly indicate that MI and TIMP

comprise a gene family (21). Conceivably, additional members of the family will be found. For example, other reported inhibitors (e.g., refs. 17, 22, and 23) could be related to TIMP and MI.

Human TIMP is a glycoprotein (6–9) and has two potential N-linked glycosylation sites (refs. 8 and 9; Fig. 2). In contrast, the bovine MI does not bind to columns of immobilized lectins (10), the unglycosylated *E. coli*-expressed human MI comigrates with the natural bovine MI in SDS/polyacrylamide gels, and there are no potential N-linked glycosylation sites in either bovine or human MI. These observations suggest that natural MI is unglycosylated, unless there is some minor O-linked glycosylation.

The existence of this family raises several questions as to potential differences in the physiological roles of the inhibitors. TIMP is known to inhibit interstitial collagenase, gelatinases, and stromelysin (6, 24–27) and to form a stoichiometric complex with collagenase (28, 29); at present, MI has been shown to inhibit interstitial collagenase and gelatinases (10). Thus there may be overlap in their interactions with metalloproteinases. As mentioned, the metalloproteinases make up a gene family themselves, and there appears to be both specificity and overlap in their degradation of extracellular matrix substrates. It will be of interest to determine whether there are any physiologically important distinctions in the interactions between metalloproteinase and inhibitor family members with regard to mechanisms, specificities, and affinities. Distinctions relevant to these interactions could also exist in tissue distribution, compartmentalization, and regulation of expression. TIMP is present in many tissues (see review, ref. 1), and expression of both TIMP and metalloproteinases can be influenced by a variety of lymphokines, growth factors, and other agents, including interleukin 1 and transforming growth factor β (refs. 30 and 31; see review, ref. 1). The availability of cDNA probes and antibodies will enable the distribution and regulation of expression of MI to be compared. In addition, availability of recombinant-derived MI makes it possible to explore therapeutic uses. Studies of the effects of TIMP on tumor cell and endothelial cell invasion of connective tissue *in vitro* (25, 32, 33) and on experimentally induced arthritis and tumor metastasis in animals (34, 35), have been reported.

While this paper was in review, the complete amino acid sequence (determined by sequencing of isolated protein) of a metalloproteinase inhibitor from human melanoma cells was reported (36). The inhibitor had 41% amino acid identity to TIMP and was designated TIMP-2. In comparison with the amino acid sequence of human aortic endothelial cell MI deduced in the present manuscript, the TIMP-2 sequence (36) differs at 10 positions and has 2 less residues. It remains to be determined whether these differences are due to errors in nucleotide or amino acid sequencing, to polymorphism, or to the existence of multiple genes. Also, partial amino acid sequences (totaling 55 residues) of an inhibitor from simian virus 40-transformed human lung fibroblasts were reported (37). These partial sequences exactly match regions from the human MI sequence deduced here and the reported human TIMP-2 (36).

We thank Arlette Harcourt, Jessica Katzowitz, and Kenneth Chen for assistance with the DNA sequencing; Tsuey-Dawn Yean for expert technical assistance; Barry Ratzkin for helpful discussions; and Julie Heuston and Joan Bennett for preparation of the figures and manuscript. This work was supported in part by Grant CA 42919 from the National Institutes of Health, Department of Health and Human Services.

1. Hay, E. D. (1981) *J. Cell Biol.* **91**, 205s–223s.
2. Werb, Z. (1989) in *Textbook of Rheumatology*, eds. Kelley, W. N., Harris, E. D., Jr., & Sledge, C. B. (Saunders, Philadelphia), 3rd Ed., pp. 300–321.
3. Perez-Tamayo, R. (1982) in *Collagen in Health and Disease*, eds. Weiss, J. B. & Jayson, M. I. (Churchill Livingstone, New York), pp. 135–159.
4. Kuettner, K. E., Hiti, J., Eisenstein, R. & Harper, E. (1976) *Biochem. Biophys. Res. Commun.* **72**, 40–46.
5. Travis, J. & Salvesen, G. L. (1983) *Annu. Rev. Biochem.* **52**, 655–709.
6. Murphy, G., Cawston, T. E. & Reynolds, J. J. (1981) *Biochem. J.* **195**, 167–170.
7. Stricklin, G. P. & Welgus, H. G. (1983) *J. Biol. Chem.* **258**, 12252–12258.
8. Docherty, A. J. P., Lyons, A., Smith, B. J., Wright, E. M., Stephens, P. E., Harris, T. J. R., Murphy, G. & Reynolds, J. J. (1985) *Nature (London)* **318**, 65–69.
9. Carmichael, D. F., Sommer, A., Thompson, R. C., Anderson, D. C., Smith, C. G., Welgus, H. G. & Striklin, G. P. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2407–2411.
10. De Clerck, Y. A., Yean, T.-D., Ratzkin, B. J., Lu, H. S. & Langley, K. E. (1989) *J. Biol. Chem.* **264**, 17445–17453.
11. Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y. & Matsubaro, K. (1985) *J. Biol. Chem.* **260**, 2605–2608.
12. Denhardt, D. T. (1966) *Biochem. Biophys. Res. Commun.* **23**, 641–646.
13. Sanger, R., Nicklen, S. & Coulson, R. A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
14. Burnette, W. N., Mar, V. L., Cieplak, W., Morris, C. F., Kaljot, K. T., Marchitto, K. S., Sachdev, R. K., Locht, C. & Keith, J. M. (1988) *Bio/Technology* **6**, 699–706.
15. Johnson-Wint, B. (1980) *Anal. Biochem.* **104**, 175–181.
16. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
17. Herron, G. S., Banda, M. J., Clark, E. J., Gavrilovic, J. & Werb, Z. (1986) *J. Biol. Chem.* **261**, 2814–2818.
18. Proudfoot, N. J. & Brownlee, G. G. (1974) *Nature (London)* **252**, 359–362.
19. Watson, M. E. E. (1984) *Nucleic Acids Res.* **12**, 5145–5164.
20. Morrissey, J. H. (1981) *Anal. Biochem.* **117**, 307–310.
21. Dayhoff, M. O., Barker, W. C. & Hunt, L. T. (1983) *Methods Enzymol.* **91**, 524–545.
22. Macartney, H. W. & Tschesche, H. (1983) *Eur. J. Biochem.* **130**, 79–83.
23. Murray, J. B., Allison, K., Sudhalter, J. & Langer, R. (1986) *J. Biol. Chem.* **261**, 4154–4159.
24. Bunning, R. A. D., Murphy, G., Kumar, S., Phillips, P. & Reynolds, J. J. (1984) *Eur. J. Biochem.* **139**, 75–80.
25. Thorgeirsson, U. P., Liotta, L. A., Kalebic, T., Margulies, I. M., Thomas, K., Rios-Candelore, M. & Russo, R. G. (1982) *J. Natl. Cancer Inst.* **69**, 1049–1054.
26. Chin, J. R., Murphy, G. & Werb, Z. (1985) *J. Biol. Chem.* **260**, 12367–12376.
27. Okada, Y., Nagase, H. & Harris, E. D., Jr. (1986) *J. Biol. Chem.* **261**, 14245–14255.
28. Cawston, T. E., Murphy, G., Mercer, E., Galloway, W. A., Hazleman, B. L. & Reynolds, J. J. (1983) *Biochem. J.* **211**, 313–318.
29. Welgus, H. G., Jeffrey, J. J., Eisen, A. Z., Roswit, W. T. & Stricklin, G. P. (1985) *Collagen Relat. Res.* **5**, 167–179.
30. Murphy, G., Reynolds, J. J. & Werb, Z. (1985) *J. Biol. Chem.* **260**, 3079–3083.
31. Edwards, D. R., Murphy, G., Reynolds, J. J., Whitham, S. E., Docherty, A. J. P., Angel, P. & Heath, J. K. (1987) *EMBO J.* **6**, 1899–1904.
32. Mignatti, P., Robbins, E. & Rifkin, D. B. (1986) *Cell* **47**, 487–498.
33. Mignatti, P., Tsuboi, R., Robbins, E. & Rifkin, D. B. (1989) *J. Cell Biol.* **108**, 671–682.
34. Schultz, R. M., Silberman, S., Persky, B., Bajkowski, A. S. & Carmichael, D. F. (1988) *Cancer Res.* **48**, 5539–5545.
35. Carmichael, D. F., Stricklin, G. P. & Stuart, J. M. (1989) *Agents Actions* **27**, 378–379.
36. Stetler-Stevenson, W. G., Krutzsch, H. C. & Liotta, L. A. (1989) *J. Biol. Chem.* **264**, 17374–17378.
37. Goldberg, G. I., Marmer, B. L., Grant, G. A., Eisen, A. Z., Wilhelm, S. & He, C. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8207–8211.