# Structural organization and chromosomal localization of the mouse collagenase type I gene

Marina SCHORPP,* Marie-Genevieve MATTEI,† Ingrid HERR,* Sabine GACK,* Jörg SCHAPER* and Peter ANGEL‡

*Forschungszentrum Karlsruhe, Institut für Genetik, Postfach 3640, 76021 Karlsruhe, Germany, and †INSERM U406, Hopital d'enfants de la Timone, 13385 Marseille, Cedex 5, France

A clone containing genomic sequences of part of the murine collagenase type 1 (MMP-1) gene was isolated. It contains exons 1–6 encoding all the domains required for collagenase function and 9 kb of 5′-flanking sequences. The gene organization and exon/intron borders are highly similar to the already described human and rabbit MMP-1 genes. However, neither the intron sequences, nor the promoter region up to position −660 exhibit significant sequence homologies with rabbit and human MMP-1, except for an AP-1-binding site and two PEA-3 consensus sequences. Binding studies *in vitro* revealed that the AP-1-binding site is recognized by Fos/Jun heterodimers with very high affinity. By *in situ* hybridization the mouse MMP-1 gene was located to the A1-A2 region of chromosome 9 in proximity to the curly whiskers (*cw*) locus. Based on the lack of sequence homologies of the promoter and intron regions, and since the chromosomal localization of the mouse and human MMP-1 genes may not be syntenic, these data strongly support previous suggestions that the MMP-1 genes from mouse, compared with rabbit and human, have evolved from different ancestoral genes. The presence of the AP-1- and PEA-3- binding sites in all mammalian MMP-1 genes isolated so far, may, however, suggest evolutionary selection for common regulatory mechanisms of MMP-1 transcription.

## INTRODUCTION

The extracellular matrix (ECM) plays an essential role during normal tissue development, function and remodeling. Since the ECM has to undergo continuous changes during processes that require tissue remodeling, such as embryogenesis, organogenesis and wound healing, the enzymes capable of degrading components of the ECM, the family of the matrix-degrading metalloproteinases (MMPs), are thought to play an instructive role during these processes (for review, see [1–5]). The expression and activity of MMPs has to be tightly controlled; inappropriate and excessive activity has been proposed to be involved in a variety of pathological processes, such as rheumatoid joint destruction, corneal ulceration, metastasis of tumour cells, and genetic diseases (e.g. recessive dystrophic epidermolysis bullosa [for review, see [1–5]).

The MMPs can be divided into three subclasses: collagenases, gelatinases and stromelysins. These share a high degree of structural similarity but differ in substrate specificity and in response to extracellular stimuli, suggesting an individual function for each member [1–5].

Collagenase type 1 (MMP-1) is primarily responsible for the turnover of the interstitial collagens type I, II and III [1–5]. In tissue culture cells, MMP-1 transcription is regulated by growth factors, mediators of inflammation, oncogenes, phorbol ester tumour promoters and carcinogens [1–6]. The different signal transduction pathways initiated by these stimuli converge at the AP-1-binding site, which also constitutes the phorbol ester-responsive element (TRE [7]). AP-1 is a dimeric complex whose subunits are encoded by the members of the *jun*, *fos* and ATF gene families (for review, see [6]). Alterations of AP-1 activity to control MMP-1 transcription are due to the rapid post-translational modification of pre-existing Fos and Jun proteins, the subsequent transcriptional activation of the *c-fos* and *c-jun* genes, and physical interaction between Fos/Jun dimers and other cellular transcription factors, such as steroid hormone receptors (for review, see [8]). Additional *cis*-acting elements in the human MMP-1 promoter interact with factors that modulate AP-1 activity, such as members of the Ets family [9], and the transforming growth factor-$\beta$ (TGF-$\beta$) inhibitory factor (TIF), which mediates repression of growth factor-induced MMP-1 expression by TGF-$\beta$ through binding to a specific recognition sequence in the promoter region [10]. In agreement with the regulatory function of AP-1 in tissue culture cells we have shown previously that in *fos*-transgenic mice, MMP-1 expression is increased in a tissue-specific manner which correlates with the sites of Fos-induced long-term phenotypic alterations [11].

In the present study we describe the chromosomal localization and isolation of part of the mouse MMP-1 gene. By *in situ* hybridization the gene was mapped to the A1-A2 region of chromosome 9 in close proximity to the curly whisker (*cw*) gene locus. Characterization of the promoter region and intron/exon sequences did not reveal significant sequence homologies, except for one consensus binding site for the transcription factors TATA-binding protein, two ets/PEA3-binding sites and a consensus AP-1 site, supporting previous suggestions that the MMP-1 gene from mouse (and possibly rat) and from other mammalian species may have evolved from different ancestoral genes [12]. The AP-1 site is recognized by Fos/Jun heterodimers with very high affinity, suggesting that the Fos-dependent tissue-specific induction of MMP-1 in *fos*-transgenic mice is a direct effect of Fos binding to the MMP-1 promoter. The presence of the PEA3-

Abbreviations used: MMP-1, collagenase type 1; *cw*, curly whiskers; ECM, extracellular matrix; MMPs, matrix-degrading metalloproteinases; TRE, phorbol ester-responsive element; TGF-$\beta$, transforming growth factor $\beta$; TIF, TGF-$\beta$ inhibitory factor; PMA, phorbol 12-myristate 13-acetate.

‡ To whom correspondence should be addressed.

and AP-1-binding sites in all mammalian MMP-1 genes isolated so far strongly suggests evolutionary selection for common mechanisms of transcriptional control of the MMP-1 genes.

## MATERIALS AND METHODS

### Cloning and sequencing of the mouse MMP-1 gene

A mouse 129/OLA genomic library (provided by A. Berns, Amsterdam, The Netherlands) was screened using a random-labelled cDNA probe for mouse MMP-1 [11]. Hybridizations and washes were performed using the Church and Gilbert protocol [13]. For subcloning and sequencing of the phage DNA inserts, various restriction fragments of the genomic clone were identified by Southern-blot analysis and subsequently cloned into pBluescript SK (Stratagene). Sequencing was performed according to the dideoxynucleotide chain termination procedure using a T7 sequencing kit (Pharmacia).

### DNA amplification using PCR

In order to determine the size of introns and exon/intron borders, PCR was performed to amplify regions between individual exons. The following oligonucleotides were used as specific primer pairs: for exons 1 and 2 (E1-E2), 5′-GCATTCAGCTATCCTGGCCAC-3′ and 5′-TGGAGGTCACTGTAGAC-3′; for E2-E3, 5′-ATGAGAAAACCAAGATGTGGAGTGCCTGATGT-3′ and 5′-CAGACCAGACCTTGAAGGC-3′; for E3-E4, 5′-GCCTTCAAGGTCTGGTCTG-3′ and 5′-TTTCATCATCATCAAAATGGGCATCCCCACCA-3′; for E4-E5, 5′-TCCTGGACCAAACTA-3′ or 5′-TGGTGGGGATGCCCATTTTGATGATGATGAAA-3′ and 5′-TCCTTGGAGTGGTCCAGACGCAGGGAGTGGCCAAGCTCATGGGCAGCAAC-3′; for E5-E6, 5′-GTTGCTGCCCATGAGCTTGGCCACTCCCTAGGTCTGGATCACTCCAAGGA-3′ and 5′-GGTCACACTTCTCTGGT-3′.

### Mapping of the start site of transcription

The genomic *Hpa*I-*Acc*I restriction fragment (−69/+292) was subcloned in pBluescript SK (Stratagene), spanning 59 nt of the second exon, the first intron, the first exon and part of the putative promoter region. The antisense probe was synthesized *in vitro* using T7 RNA polymerase and hybridized with 5 μg of poly(A)$^+$ RNA from untreated and phorbol 12-myristate 13-acetate (PMA)-treated (6 h) mouse embryonic fibroblasts. The analysis was performed as described previously [14], except that the hybridization was carried out at 45 °C. To determine the exact size of the exons, the protected fragments were run on a 6% sequencing gel along with $^{35}$S- and $^{32}$P-labelled sequencing reactions of the identical region of the MMP-1 gene. The oligonucleotide 5′-TGCAAACACAAGGTCTTCCTC-3′ used as primer encodes the 3′ part of the first exon, except for the last three nucleotides.

### Cell culture and RNA isolation

Primary mouse embryonic fibroblasts were grown in Dulbecco's modified Eagle's Medium supplemented with 10% fetal calf serum. Cells were treated for 8 h with 60 ng/ml PMA. Total RNA was prepared according to Chomczynski and Sacchi [15].

### Gel retardation analysis

Nuclear extracts from NIH3T3 and HeLa TK$^-$ cells were prepared as described by Dignam et al. [16]. Gel retardation assays were performed as described previously [17,18]. The following double-stranded oligonucleotides were used as probes: 5′-AGTGGTGACTCATCACT-3′ and 5′- AGTGATGAGTCACCACA-3′ [mouse (m)-col AP-1, position −52 to −36]; 5′-AGCTAGCATGAGTCAGACAC-3′ and 5′-AGCTGTGTCTGACTCATGCT-3′ [human (h)-col AP-1, position −76 to −61] and 5′-AGCTAGCATTACCTCATCCC-3′ and 5′-GATCGGGATGAGGTAATGCT-3′ (jun2: *c-jun*, position −194 to −179). The rabbit polyclonal anti-human Fos and Jun antisera have been described previously [17,18].

### Chromosomal localization

*In situ* hybridization experiments were carried out as described previously [19] using metaphase spreads from a WMP male mouse in which all the autosomes except 19 were in the form of metacentric robertsonian translocations. Concanavalin A-stimulated lymphocytes were cultured at 37 °C for 72 h with 5-bromodeoxyuridine added for the final 6 h of culture (60 mg/ml of medium) to ensure a chromosomal R-banding of good quality. As a probe, an MMP-1 cDNA-fragment of 423 bp, encompassing nucleotides 731–1153 of the coding sequence, was amplified by PCR using the oligonucleotides 5′-CAAGGACCCAGGAGCCCTGAT-3′ and 5′-CCCAGGTCAGATATTTTTCTGGG-3′ as specific primers and subcloned as an *Eco*RV-*Pst*I fragment into a pSP65-derived vector (pAZ; U. Günthert, unpublished work). This cDNA fragment was $^3$H-labelled by nick-translation to a specific activity of $1.6 \times 10^6$ d.p.m./mg. The radiolabelled probe was hybridized to metaphase spreads at a final concentration of 25 ng/ml in hybridization solution as previously described [19]. After coating with nuclear track emulsion (Kodak NTB2), the slides were exposed for 26 days at +4 °C, then developed. To avoid any slipping of silver grains during the banding procedure, chromosome spreads were first stained with buffered giemsa solution and metaphases photographed. R-banding was then performed by the fluorochrome–photolysis–giemsa method and the metaphases were rephotographed before analysis.

## RESULTS

### Structural organization of the MMP-1 gene

Using a mouse MMP-1 cDNA probe encoding the putative $Zn^{2+}$-binding domain of the enzyme [11], we were able to isolate a genomic clone (containing 13 kb of genomic sequences) from a mouse 129 library. Subcloning and sequencing of the clone revealed that it encodes exons 1–6, and approximately 9 kb of 5′ flanking sequences. The structural organization of the isolated gene fragment is schematically illustrated in Figure 1. As described below, exons 1–6 encode the domains that are required for secretion, activation and enzymic activity of the MMP-1 protein (for review, see [1–5]). Rescreening of the library did not give rise to the identification of additional clones encoding the C-terminus of the MMP-1 protein and 3′-untranslated sequences of the MMP-1 mRNA.

Using synthetic oligonucleotide primers derived from the MMP-1 cDNA, the nucleotide sequence of the exons and flanking intron sequences was determined (Table 1). Exon/intron boundaries were defined by the termination of colinearity with the cDNA sequences. The sizes of the introns 1–6 were determined either by complete sequencing (introns 1 and 3), or by PCR amplification of intron regions using specific primer pairs derived from flanking exon sequences. The analysis revealed that the size of exons 1–6 varies from 118 to 242 bp, whereas the introns range between 80 bp and 2.2 kb. All exon/intron junctions were found to contain splice donor (AG) and acceptor (GT) sites with consensus sequences for RNA splicing ([20]; Table 1). The exon

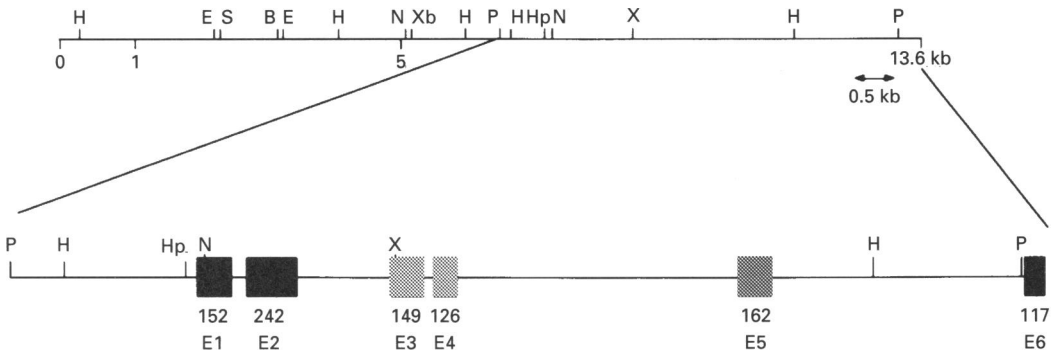


**Figure 1 Structural organization of part of the MMP-1 gene**

Top, a partial restriction map covering 13.6 kb of the mouse MMP-1 locus is shown. The following restriction sites are displayed: H (*Hin*dIII), E (*Eco*RI), S (*Sac*I), N (*Nsi*I), Xb (*Xba*I), P (*Pst*I), Hp (*Hpa*I) and X (*Xho*I). Bottom, an enlarged diagram of the gene with exons depicted as boxes is outlined. The sizes of the individual exons and introns (in nucleotides) are indicated.

**Table 1 Exon-intron structure of the mouse collagenase type 1 gene**

Nucleotide sequences at the intron and exon junctions are shown. The derived amino acid sequence is displayed below. Amino acids encoded by split codons are listed and are given in parentheses.

| Exon/Intron Number | Exon/Intron Junctions | Exon Size (bp) | Intron Size (bp) |
|---|---|---|---|
| 1 29bp untransl. | ATG CAT TCA ... ... TTT GCA GAG gtactgtatt<br>Met His Ser Phe Ala Glu | 152 | 86 |
| 2 tcttttctag | CAC TAC TTG ... ... TTA ACT TAC AG gtaaagcaga<br>His Tyr Leu Leu Thr Tyr Ar(g) | 242 | ≈620 |
| 3 tgttgaatag G | ATT GTG AAC ... ... GGG ACT AAA G gtaagggtgt<br>(Ar)g Ile Val Asn Gly Thr Lys G(lu) | 149 | 167 |
| 4 gtctgcatag AA | CAT GGT GAC ... ... AGT TCC AAA G gtaatcttcc<br>(G)lu His Gly Asp Ser Ser Lys G(ly) | 126 | ≈1500 |
| 5 tttcttacag GC | TAC AAC TTG ... ... TTT CTT TAT G gtaaacttaa<br>(Gl)y Tyr Asn Leu Phe Leu Tyr G(ly) | 162 | ≈1300 |
| 6 gtcactgcag GT | CCA GGC GAT ... ... TTT AAA GAC AG gtattcttaa<br>(Gl)y Pro Gly Asp Phe Lys Asp Ar(g) | 117 | – |

sequences were found to be identical with the sequences of the previously described mouse MMP-1 cDNAs [11,12]. Exon 1 encodes the signal peptide (residues 1–23) and the first 18 residues of the propeptide (residues 24–41). Exon 2 encodes amino acids 42–122, representing the propeptide including the cystein switch region (residues 42–104), and part of the catalytic domain. Exons 3–5 carry residues 123–171, 172–214 and 215–268 respectively, which represent the catalytic domain (105–268) including the zinc binding domain (220–234) which is encoded by exon 5. Finally, exon 6 (269–307) encodes the hinge region and part of the hemopexin-like domain (residues 269–472 of the complete protein).

Despite the high degree of homology in the overall modular structure of the coding sequences, neither the sizes nor the sequences of the introns exhibited significant homology to the human [21] and rabbit [22] MMP-1 genes.

## Characterization of the promoter region of the mouse MMP-1 gene

To determine the start site of transcription we performed ribonuclease protection analyses of poly(A)$^+$ RNA from untreated and PMA-treated (8 h) mouse embryonic fibroblasts, which are known to express interstitial collagenase [11]. The antisense riboprobe encompassed 59 nucleotides of the second exon, the first intron, the first exon and part of the putative promoter region. Both in RNA from untreated and PMA-treated cells, two specific fragments were protected from ribonuclease digestion by the antisense probe (Figure 2a). The exact size of the fragments was determined by running $^{35}$S- and $^{32}$P-labelled sequencing reactions from the same region of the genomic clone along with the protected fragments. Based on these results the sizes of the fragments encoding the first and part of the second exon are 152 bp and 59 bp respectively. Using a larger riboprobe which also protected the complete exon 2, two fragments of 152 and 242 bp became visible (results not shown). Based on these results, the previously isolated mouse MMP-1 cDNA [12] lacks 19 nucleotides of the 5′-untranslated sequences.

Analysis of the 5′ flanking sequences (Figure 2b) revealed the presence of various consensus recognition sites for DNA-binding proteins, such as a consensus TATA motif, TATAAAA, at position −28 to −22, an AP-1 consensus sequence (TGACTCA) at position −47 to −41 and the sequences AGGAAGT (position −74 to −68) and AGGAAGA (position −121 to −115), which resemble the consensus sequence of a PEA3-binding site [9]. In addition, a putative binding site for TIF was found at position −128 to −137 (GGTTTGTGGC) which, however, contains one mismatch compared with the consensus TIF-binding site

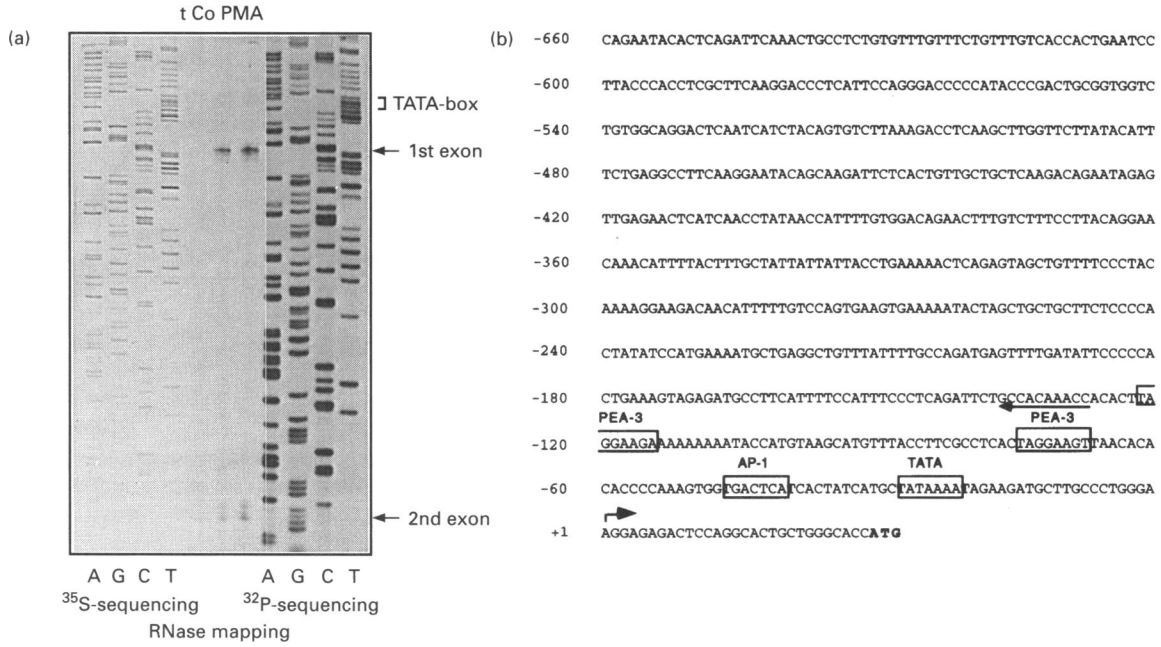


**Figure 2 The mouse MMP-1 promoter**

(**a**) Determination of the start site of transcription by RNAse protection analysis. The mRNA start site was determined by ribonuclease mapping using an anti-sense RNA probe (−65 to +292) encompassing part of the second exon, the first intron, the first exon and part of the putative promoter region. This riboprobe was hybridized to poly(A)$^+$ RNA from untreated (Co) and PMA-treated (PMA, 6 h) embryonic fibroblasts known to express MMP-1. As a negative control, t-RNA (t) was used. Upon RNase digestion, a fragment of 59 nt representing the second exon is visible. This is in agreement with the expected size of the protected fragment representing part of the second exon based on DNA sequencing (Table 1). In addition, a larger fragment of 152 nt, representing exon 1, became resistant to RNAse digestion. The exact size of the RNAse resistant fragments was measured by co-running $^{35}$S- and $^{32}$P-sequencing reactions from the same region of cloned genomic DNA. Note that the sequence of the primer used for the sequencing reactions is missing the last three nucleotides of exon 1. This has been taken into account when the size of the first exon was determined as 152 nt. (**b**) The promoter of the mouse MMP-1 gene contains conserved transcription factor binding sites. The nucleotide sequence at the 5′-end of the mouse MPP-1 gene, including the coding sequence up to the first amino acid codon, is given. The numbering of nucleotides starts at the transcription initiation site. Conserved putative transcription factor binding sites are boxed. The putative TIF-binding site (one mismatch compared with the consensus sequence, opposite orientation) is underlined.

(GNNTTGGtGa; [10]). No other consensus binding sites of known transcription factors were found between positions −140 and −660.

## cFos/cJun heterodimers bind to the mouse MMP-1 AP-1-binding site with very high affinity

To confirm that the putative AP-1 consensus sequence at position −47 to −41 is indeed recognized by Fos/Jun proteins we performed binding studies *in vitro*. A 17 bp oligonucleotide probe containing the mouse AP-1 motif, mouse-col AP-1, was incubated with nuclear extract from mouse NIH3T3 cells. One specific protein–DNA complex could be observed that can be competed, at least partially, by a 10-fold excess of non-labelled probe (Figure 3a, left-hand side, lanes 1 and 2). Binding is completely lost using a 100-fold excess of non-labelled probe (lane 4). Interestingly, a 100-fold excess of the 17 bp human collagenase TRE, human-col AP-1, is needed to get efficient competition (lane 7), suggesting that the AP-1 site of the mouse promoter has a higher affinity compared with its human counterpart. This interpretation was confirmed by the reverse experiment (Figure 3a, right-hand side), in which an oligonucleotide probe containing the human TRE, with a 10-fold excess of the non-labelled mouse TRE probe, was sufficient to block binding to the human TRE completely. A mutated version of the AP-1 consensus sequence [7] did not affect complex formation (results not shown).

In summary, the mouse TRE has an at least 10-fold higher affinity for AP-1 binding compared with the human sequence. One possible explanation for these differences may be the flanking sequences surrounding the AP-1 consensus binding sites, as, in the mouse, but not in the human collagenase promoter, the palindromic structure of the consensus sequence is extended by five nucleotides (Figure 3a).

To confirm the presence of Fos and Jun proteins in the complex binding to the mouse-col AP-1 site, we used antibodies that specifically recognize all members of either the Fos or Jun protein family. The electrophoretic mobility of the protein complex binding to the DNA was completely altered in the presence of a polyclonal antibody directed against the Fos proteins since a complex with slower mobility (supershift) is detected (Figure 3c). This complex formation is specific, since in the presence of pre-immune serum, DNA-binding is not affected.

We also used a polyclonal antibody directed against the highly conserved DNA-binding domain of the Jun proteins [17,18] which interferes with DNA binding *in vitro* of complexes containing cJun, JunB or JunD to the human coll-TRE (H. van Dam and P. Angel, unpublished work). In the presence of Jun-specific antibodies, complex formation with the mouse collagenase TRE is completely lost (Figure 3c).

The antisera recognizing either Fos or Jun proteins yielded completely altered protein–DNA complexes, generating either supershifted complexes (αFos), or forming protein–antibody complexes that cannot bind to the mouse TRE (αJun). Therefore, the heterodimeric subunits of the specific complex have to consist of members of the Fos and Jun protein families. As a negative

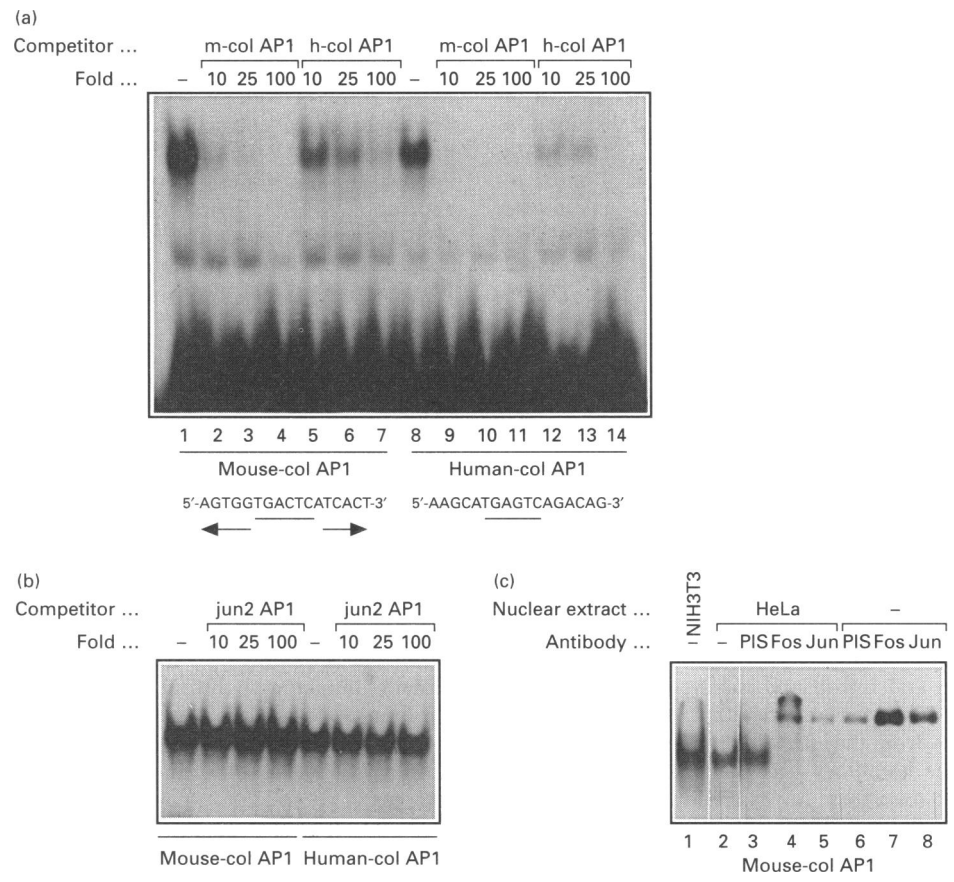


**Figure 3 Fos/Jun binds with very high affinity to the mouse MMP-1 AP-1 site *in vitro***

(**a**) Nuclear extract (5 $\mu$g) of mouse NIH3T3 cells, which were treated with PMA for 6 h before extraction, was incubated with radioactively labelled probes of the mouse or human collagenase AP-1 sites for 30 min at room temperature. In parallel, a 10, 25 or 100-fold molar excess of either an unlabelled mouse or human MMP-1 AP-1 oligonucleotide probe was added as competitor, as indicated at the top of the Figure. The DNA–protein complexes were resolved on 5% non-denaturing PAGE. The migration of specific complexes and free unretarded probe are shown. The sequences of the mouse and human oligonucleotides used are shown at the bottom of the Figure. The consensus TRE sequence is underlined. The arrows mark the extended palindromic nature of the mouse AP-1-binding site. (**b**) Gel retardation assay was performed as described in (**a**). As competitor, an oligonucleotide of the jun2 AP-1 site, 5'-AGCTAGCATTACCTCATCCC-3', was used, which is specifically recognized by cJun/ATF-2 heterodimers. (**c**) Nuclear extracts (5 $\mu$g) from untreated HeLa cells were incubated on ice for 2 h with a polyclonal antiserum directed against Fos or Jun proteins, or with pre-immune rabbit serum. Subsequently, the radioactively labelled mouse MMP-1 AP-1 oligonucleotide was added and incubated at room temperature for 30 min (lanes 3, 4 and 5). As a control, 5 $\mu$g of nuclear extracts from NIH3T3 or HeLa cells was incubated in the absence of antiserum (lanes 1 and 2). Note that the incubation of the antiserum or the pre-immune serum in the absence of extract gave rise to a non-specific complex (lanes 6, 7 and 8).

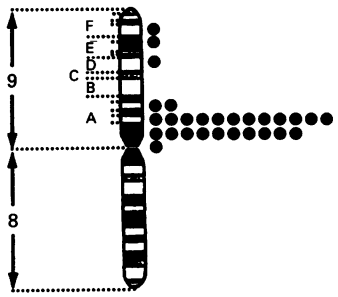


**Figure 4 Localization of the MMP-1 gene to mouse chromosome 9 by *in situ* hybridization**

Diagram of WMP mouse RB (8,9) chromosome, indicating the distribution of labelled sites. A–F refer to the idiogram of the banding patterns in mouse chromosomes revealed by Giemsa staining.

control we used a variant AP-1 site of the human *c-jun* promoter containing an additional nucleotide in the centre of the consensus sequence (jun2, TTACCTCA) as competitor. This sequence was previously identified to bind predominately cJun/ATF-2 heterodimers but is not recognized by Fos/Jun heterodimers [17,18]. This site cannot compete for binding to the human collagenase TRE nor for binding to the mouse collagenase TRE (Figure 3b).

These results demonstrate that the AP-1 site of the mouse MMP-1 is recognized by Fos/Jun heterodimers, but not by other members of the AP-1 superfamily, such as ATF-2.

## Chromosomal localization of the mouse MMP-1 gene

To determine the chromosomal location of the MMP-1 gene, metaphase spreads from a WMP male mouse were hybridized *in situ* with a 422 bp mouse MMP-1 cDNA probe (Figure 4). The hybridization probe encodes nucleotides 668–1153 of the coding

sequence which is unique for the MMP-1 sequence, but does not cross-hybridize with any other member of the MMP family. In 100 metaphase cells examined after *in situ* hybridization, there were 156 silver grains associated with chromosomes and 28 of these (17.9%) were located on chromosome 9. The distribution of grains on this chromosome was not random: 22 out of the 28 silver grains (78.6%) mapped to the A1-A2 region of chromosome 9. These results allow us to map the MMP-1 gene to the A1-A2 region of chromosome 9 of the mouse genome. Within this region, the *cw* locus has been localized [23,24]. Additional mouse mutants have not yet been mapped to this region [23,24].

## DISCUSSION

The protein family of MMPs consists of at least nine members which share a high degree of structural similarity but differ in substrate specificity, suggesting specific functions of individual MMPs in the degradation of components of the extracellular matrix. This paper describes the isolation and characterization of genomic sequences of part of the mouse MMP-1 gene encoding exons 1–6 and approximately 9 kb of 5′-flanking sequences and the determination of the chromosomal location of the MMP-1 gene.

Sequence comparison of the promoter region (up to position −660), the first exon and intron and the second exon to the human and rabbit MMP-1 genes revealed that the mouse MMP-1 gene differs markedly from the interstitial collagenases of other species (Figure 5). The homology between human and murine MMP-1 is even lower than that, for example, between human neutrophile and interstitial collagenase [3]. While the 5′-untranslated sequences and first intron do not exhibit any sequence homologies (except for the splice donor and splice acceptor consensus sequences), low homology is found for the first exon encoding the signal peptide and part of the propeptide. Homology is significantly increased in the second exon encoding the highly conserved 'cystein switch' region required for enzyme activation [3], suggesting evolutionary selection for a specific protein structure required for enzyme function. The divergence of mouse MMP-1 from other mammalian MMP-1 genes strongly supports the previous hypothesis (based on phylogenetic analysis of cDNA sequences), that the order Rodentia is an out-group to the other eutherian (placental) mammalian orders [12].

Despite the overall lack of sequence homology of the promoter regions of the mouse and human (and rabbit) MMP-1 genes, three *cis*-acting elements are present that are commonly found in all mammalian MMP-1 genes isolated so far, suggesting evolutionary selection for common routes of transcriptional regulation of the MMP-1 genes. These sites include a TATA element 28 bp upstream from the transcriptional start site, two ets/PEA-3 consensus sequences [9] and an AP-1-binding site [7]. The presence of the AP-1- and ets/PEA-3-binding sites is of particular interest since both sites synergize in the regulation of the human and rabbit MMP-1 genes and other matrix metalloproteinases, such as stromelysin-1, in response to growth factors, cytokines, carcinogens, tumour promoters and certain oncogenes [3,6,9]. While the AP-1 site was found to confer inducibility in the absence of the PEA-3 site [7,9,21], the PEA-3 site is not active on its own, but strongly augments AP-1-dependent induction of the MMP-1 promoter, depending on the appropriate spacing between the AP-1 and PEA-3 sites ([9,21]. However, since the mouse MMP-1 promoter contains 11 extra nucleotides between the PEA-3 and AP-1 sites, synergistic action of AP-1 and Ets proteins (binding to the proximal and distal PEA-3 sites) has yet to be confirmed. It is possible that, due to the very high affinity of the AP-1 site, DNA binding and transcriptional activation by Fos/Jun heterodimers may not depend on the Ets proteins. In addition, the spacing between the AP-1 site and the TATA box is reduced in the murine promoter compared with the human and rabbit genes (Figure 5). Therefore, other factors such as the protein interacting with footprint 'box-1' of the human MMP-1 promoter (located beween AP-1 and TATA sites; [7]) cannot interact with the mouse MMP-1 promoter, which may affect the efficiency of formation of the RNA polymerase II–preinitiation complex.

The presence of the high-affinity binding site strongly suggests that the previously observed enhanced expression of MMP-1 in *c-fos* transgenic mice [11] is a direct effect of c-Fos binding to the MMP-1 promoter. It is important to note however, that, despite vast overexpression of c-Fos in almost every tissue, enhanced levels of MMP-1 transcripts are observed in only a limited number of tissues [11], suggesting an important role of additional factors (possibly expressed in a tissue-specific manner) that cooperate with AP-1 to regulate MMP-1 expression in the *c-fos* transgenic mice.

The murine MMP-1 gene is localized on the A1-A2 region of chromosome 9. Previously, a cluster of human metalloproteinase genes including MMP-1, stromelysin-1 (MMP-3) and stromelysin-2 (MMP-10) was localized to chromosome 11q22-23 [25,26]. Most of the A1-A2 region of mouse chromosome 9 corresponds to human chromosome 19 [24]. This lack of obvious syntenic

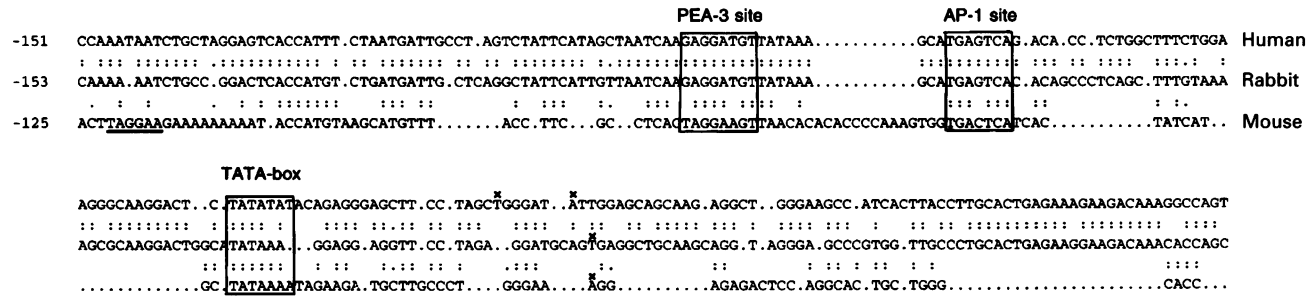


**Figure 5 Sequence comparison and alignment of the nucleotide sequence of the promoter and part of the first exon in the human, rabbit and mouse MMP-1 gene**

Double points indicate conserved nucleotides between MMP-1 sequences from human and rabbit, or rabbit and mouse, or all three species, respectively. A single point indicates matches between the human and mouse genes for 2 h. The asterisks mark the transcription start sites within the individual genes. The TATA-motif and the PEA-3 and AP-1 transcription factor binding sites are boxed. The distal PEA-3 site of the mouse collagenase promoter, which is not present in the human and rabbit counterparts, is underlined.

conservation further supports the hypothesis that the mouse and human MMP-1 genes are derived from different ancestoral genes. However, there is also genetic evidence that the A2-A4 region of murine chromosome 9 has homologies with the human 11q22-q24 region [24]. Hexa, genetically mapped at 29 cM, is physically mapped to the middle of the A band of murine chromosome 9 [24] in close proximity to Ncam (28 cM). The physical mapping of Ncam is not known, but it is probably in the A band of murine chromosome 9. Interestingly, human Ncam is located on chromosome 11q23-24, suggesting that part of the 9A band of the murine genome, probably 9A2-9A4, has homologies with the human 11q22-q24 region. Since the *in situ* mapping has some degree of imprecision, synteny conservation between human and murine MMP-1 cannot be rigorously excluded. Because stromelysin-1 and stromelysin-2 colocalize with MMP-1 on human chromosome 11q23 [26], chromosomal mapping of the murine stromelysin-1 and stromelysin-2 genes may give conclusive evidence about syntenic conservation between the human and murine MMP-1 genes.

The only clearly defined genetic mouse mutation that colocalizes with MMP-1 on chromosome 9, A1-A2 [24], is *cw*. The *cw* phenotype is characterized by curled or bent whiskers [23]. In addition, a variant of the *cw* mutant, $cw^{thd}$ (tail hair depletion), exhibits abnormal or depleted whiskers and tail hair beginning at one to two weeks of age. Interestingly, these mice accept tail skin grafts from normal mice, but normal mice reject tail skin grafts from mutant mice [23]. Taking into account that whiskers are composed predominately of collagen type 1 and that MMP-1 has been proposed to play a major role during inflammation and wound healing it will be interesting to determine MMP-1 expression and enzyme activity in both *cw* mutant mouse strains.

The A1-A2 region of chromosome 9 has not been found on the basis of genetic analysis to encode genetically defined diseases that are proposed to involve MMP activity, such as alterations in organogenesis, wound healing or chronic inflammation. The lack of clearly defined mouse mutants of this region may even suggest that extensive alterations in MMP-1 expression or activity (caused by chromosomal deletion or translocation) interfere with embryonic viability. The isolation and characterization of the mouse gene described here will allow the determination of the physiological consequences of MMP-1 deficiency in the mouse system, *in vivo*, to be investigated, for example by using the powerful technique of gene targeting.

We are grateful to Anton Berns (Amsterdam, The Netherlands) for providing the mouse 129/OLA genomic library, Hans Jobst Rahmsdorf (Karlsruhe, Germany) for Jun and Fos-specific polyclonal antibodies, Christina Kelke, Matthias Becker and Ralph Müller (Karlsruhe, Germany) for technical assistance and Jonathan Sleeman (Karlsruhe, Germany) for critical reading of the manuscript. The work was supported by the ARC (Association pour la Recherche contre le Cancer) and the Ligue Nationale contre le Cancer (M.-G. M.) and by a grant from the Deutsche Forschungsgemeinschaft (An 182/6-2).

Received 19 September 1994/23 December 1994; accepted 18 January 1995

## REFERENCES

1 Alexander, C. M. and Werb, Z. (1989) Curr. Opin. Cell Biol. **1**, 974–982
2 Matrisian, L. M. (1990) Trends Genet. **6**, 121–125
3 Birkedal-Hansen, H., Moore, W. G. I., Bodden, M. K., Windsor, L. J., Birkedal-Hansen, B., DeCarlo, A. and Engler, J. A. (1993) Crit. Rev. Oral Biol. Med. **4**, 197–250
4 Liotta, L. A., Steeg, P. S. and Stetler-Stevenson, W. G. (1991) Cell **64**, 327–336
5 Murphy, G. and Reynolds, J. J. (1993) in Connective Tissue and its Heritable Disorders (Royce, P. M. and Steinmann, B., eds.), pp. 287–316, Wiley-Liss, New York
6 Angel, P. and Karin, M. (1991) Biochim. Biophys. Acta **1072**, 129–157
7 Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R., Rahmsdorf, H. J., Jonat, C., Herrlich, P. and Karin, M. (1987) Cell **49**, 729–739
8 Ponta, H., Cato, A. B. C. and Herrlich, P. (1992) Biochim. Biophys. Acta **1129**, 255–261
9 Gutman, A. and Wasylyk, B. (1990) EMBO J. **9**, 2241–2246
10 Kerr, L. D., Holt, J. T. and Matrisian, L. M. (1988) Science **242**, 1424–1427
11 Gack, S., Vallon, R., Schaper, J., Rüther, U. and Angel, P. (1994) J. Biol. Chem. **269**, 10363–10369
12 Henriet, P., Rousseau, G. G. and Eeckhout, Y. (1992) FEBS Lett. **310**, 175–178
13 Church, G. M. and Gilbert, W. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 1991–1995
14 Schorpp, M., Kugler, W., Wagner, U. and Ryffel, G. U. (1988) J. Mol. Biol. **202**, 307–320
15 Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. **162**, 156–159
16 Dignam, J. D., Lebovitz, R. M. and Roeder, R. G. (1983) Nucleic Acids Res. **11**, 1475–1489
17 van Dam, H., Duyndam, M., Rottier, R., Bosch, A., de Vries-Smits, L., Herrlich, P., Zantema, A., Angel, P. and van der Eb, A. (1993) EMBO J. **12**, 479–487
18 Herr, I., van Dam, H. and Angel, P. (1994) Carcinogenesis **15**, 1105–1113
19 Mattei, M.-G., Philip, N., Passage, E., Moisan, J. P., Mandel, J. L. and Mattei, J. F. (1985) Hum. Genet. **69**, 268–271
20 Mount, S. M. (1982) Nucleic Acids Res. **10**, 459–472
21 Angel, P., Baumann, I., Stein, B., Delius, H., Rahmsdorf, H. J. and Herrlich, P. (1987) Mol. Cell. Biol. **7**, 2256–2266
22 Fini, M. E., Plucinska, I. M., Mayer, A. S., Gross, R. H. and Brinckerhoff, C. E. (1987) Biochemistry **26**, 6156–6165
23 Green, M. C. (1989) in Genetic Variants and Strains of the Laboratory Mouse (Lyon, M. F. and Searle, A. G., eds.), pp. 12–403, Oxford University Press, Oxford
24 Lyon, M. F. and Kirby, M. C. (1993) Mouse Genome **91**, 57–58
25 Spurr, N. K., Gough, A. C., Gosden, J., Rout, D., Porteous, D. J., van Heyningen and Docherty, A. J. (1988) Genomics **2**, 119–127
26 Formstone, C. J., Byrd, P. J., Ambrose, H. J., Riley, J. H., Hernandez, D., McConville, C. M. and Taylor, A. M. (1993) Genomics **16**, 289–291