The Human Tumor Cell-derived Collagenase Stimulatory Factor (Renamed EMMPRIN) Is a Member of the Immunoglobulin Superfamily

Chitra Biswas, Ying Zhang, Rosana DeCastro, Huiming Guo, Yoshiha Nakamura, Hiroaki Kataoka, and Kazuki Nabeshima

Department of Anatomy and Cellular Biology, Tufts University School of Medicine, Boston, Massachusetts 02111

ABSTRACT

Tumor cell-derived collagenase stimulatory factor, renamed extracellular matrix metalloproteinase inducer (EMMPRIN), is a Mr. ~58,000 glycoprotein which is located on the outer surface of human tumor cells and which interacts with fibroblasts to stimulate expression of several matrix metalloproteinases in the fibroblasts. In this study, we have used several approaches to isolate a complementary DNA encoding EMMPRIN. Several peptide sequences obtained from the isolated Mr. 58,000 glycoprotein are found in the translated complementary DNA clone, verifying its identity. Computer database searches indicate that EMMPRIN is a member of the immunoglobulin superfamily and that the deduced amino acid sequence of EMMPRIN is identical to that recently reported for human basigin and M6 antigen, molecules of previously undetermined biological function.

INTRODUCTION

Degradation of extracellular matrix components of the basement membrane and interstitial matrix by MMPs is a crucial step in tumor cell invasion and metastasis (1–3). The role of tumor cell-fibroblast interactions in regulation of MMP levels in neoplasms has been demonstrated by several investigators, including ourselves (4–7). The recent finding (8–11) that, in vivo, some tumor-associated MMPs are mainly synthesized in peritumoral fibroblasts, rather than in tumor cells themselves, is consistent with a major role for these interactions in tumorigenesis in vivo.

We have shown that tumor cells in culture stimulate fibroblasts to produce high levels of collagenase and that a factor (previously termed TCSF) that is associated with tumor cell membranes, but also released into medium conditioned by tumor cells, is responsible for this stimulation (6, 12, 13). We have immunoaffinity purified the Mr. ~58,000 TCSF from a human lung carcinoma cell line, LX-1 (13, 14), and demonstrated that addition of this purified factor to cultured fibroblasts stimulates expression, not only of interstitial collagenase (MMP-1), but also of fibroblast-derived stromelysin-1 (MMP-3) and Mr. 72,000 gelatinase (MMP-2) (15, 16). While immunohistochemical studies have shown that TCSF is highly enriched around the outer surface of tumor cells and absent from most normal cells in vivo (17), recent studies from our laboratory have shown that TCSF is also present on the surface of normal human keratinocytes, where it presumably plays a role in regulating stromal MMPs (18). For these reasons, we have now renamed this factor EMMPRIN to indicate its role in extracellular matrix metalloproteinase induction via normal, as well as pathological, cellular interactions.

To help understand the chemical and biological nature of EMMPRIN and its relationship to other proteins, we have attempted to isolate cDNA clones for the protein. Oligonucleotide primers derived from peptide sequences were used to isolate EMMPRIN cDNAs by RT-PCR. Analysis of the cDNA-derived amino acid sequence of EMMPRIN indicates that it is a member of the immunoglobulin superfamily. Interestingly, the sequence is identical to that recently reported human cDNAs of unknown function, i.e., human basigin (19) and M6 antigen (20). Thus, our studies provide one important function for these proteins, namely intercellular stimulation of MMP synthesis.

MATERIALS AND METHODS

Amino Acid Sequencing. Previously, we have reported the amino acid sequences for the NH2-terminus of EMMPRIN and four peptides derived from EMMPRIN after trypsin digestion (14). We have now sequenced two more peptides derived from EMMPRIN in the same manner as described previously (14). Briefly, immunopurified EMMPRIN was subjected to SDS-PAGE and blotted to a nitrocellulose membrane. The EMMPRIN band was revealed by staining with Ponceau S. After destaining, the protein band was cut from the membrane and digested with trypsin at a ratio of 1:20 (w/w). The peptides were separated by reverse phase HPLC, and the samples were subjected to automated Edman degradation.

cDNA Synthesis. RNA was prepared from LX-1 cells by a routine procedure using guanidinium thiocyanate (21), and poly(A)+RNA was isolated using the Mini Ribosep mRNA isolation kit according to the manufacturer’s instructions (Collaborative Biomedical Products, Bedford, MA). First-strand cDNA was synthesized by reverse transcription of 1 μg of poly(A)+RNA using Moloney murine leukemia virus reverse transcriptase, according to the manufacturer’s instructions (GIBCO-BRL, Gaithersburg, MD) in the presence of random hexamers or specific primers, depending on the desired reaction product. The resulting reaction mixture was digested with RNase H and used as a template for PCR.

DNA Amplification. PCR-amplified DNA fragments were generated with a Perkin-Elmer Cetus DNA thermal cycler (Norwalk, CT) using a gene amplification kit (Perkin Elmer) according to the manufacturer’s instructions. Briefly, 100 μl of reaction mixture contained 100 ng of cDNA pool, 10 μl of 10X PCR buffer (provided in the kit), 16 μl of each deoxynucleotide triphosphate at 1.25 mm, 5 μl each of 20 μm primers, and 0.5 μl of Taq-DNA polymerase (2.5 units per assay). Samples were subjected to 30 cycles at the following conditions: 1 min at 94°C for denaturation; 1 min at 48°C for annealing; and 1.5 min at 72°C for elongation. A final elongation step consisted of a 72°C incubation for 10 min. Amplified products were separated by agarose gel electrophoresis and were identified by ethidium bromide staining. Where repamification of a PCR product was necessary, samples were applied to low-melting point agarose gels (FM, Rockland, ME); bands were excised, melted, and purified on Spin Band extraction cartridges (FMC).
Generation of an Authentic cDNA for EMMPRIN Peptide #51. Two 17-mer, degenerate, oligonucleotide mixtures, based on the previously obtained amino acid sequence of EMMPRIN peptide #51 (14), were synthesized. One of these, termed A, was synthesized in the sense direction corresponding to amino acids 2–7; the other, termed C’, was synthesized in the antisense direction corresponding to amino acids 13–18. The sequences of the oligonucleotide mixtures are shown in Table 1. These mixtures were used as primers in RT-PCR to generate an authentic EMMPRIN cDNA corresponding to amino acid residues 2–18 of peptide #51, using LX-1 cell mRNA as the initial template. The primers had 8-bp adapters with restriction site sequences (EcoR I for the 5’-end of A and Pst I for the 5’-end of C’; not shown in Table 1), allowing subsequent cloning of the product in the event that difficulty was encountered in directly ligating the product into the pCRII vector. The resulting PCR products were sized by electrophoresis against known DNA markers in 6% agarose (NuSieve GTG low melting point agarose). Several products of different intensities were identified, including the expected 66-bp product corresponding to #51 (50-bp fragment plus 2 × 8 bp adapters at 5’ and 3’ ends). The band corresponding to this size was cut out from the agarose gel and reapplied using the same primers A and C’ as described above. After checking the size of the reamplified product, all of which was 66 bp, the reaction mixture was used directly for ligating into the pCRII vector system (Invitrogen). Recombinants were selected as white colonies on plates containing 5-bromo-4-chloro-3-indolyl-b-D-galactoside. Plasmid DNA was isolated from seven of these colonies, and the sizes of the inserts were analyzed by PCR amplification using M13 primers (forward and reverse), followed by agarose gel electrophoresis.

All seven recombinants had inserts of the expected sizes. DNA from two of these clones was sequenced by the dideoxy-mediated chain termination method (22) using a double-stranded DNA cycle sequencing kit (GIBCO-BRL). The nucleotide sequence of the inserts from both clones exactly matched the amino acid sequence of peptide #51 (Fig. 1). However, the nucleotide sequence varied at three positions, all within the primer sequences, as shown in bold in Fig. 1. This suggests that several of the degenerate primers were used by the template cDNA during amplification. The use of degenerate oligonucleotide mixtures as primers often leads to variations of this kind in the regions of the cDNA that correspond to the primers (23). To avoid these regions of variation, we designed primers (B/B’) based on the sequence of the central part of the cDNA (Fig. 1) for use in the overlap extension reactions described below.

Generation of cDNAs Corresponding to the 5’ and 3’ Regions of EMMPRIN. To obtain a cDNA that includes the region of EMMPRIN that is 5’ to peptide #51, primer B’ (designed as described above) was used in the PCR in combination with primer D, a degenerate mixture corresponding to part of the previously sequenced NH2-terminal peptide #59, derived from EMMPRIN (Ref. 14, see Table 1). The PCR products were directly ligated into pCRII (Invitrogen) and used to transform Escherichia coli. Insert-containing white colonies were selected, and the plasmid DNAs were isolated, sized, and sequenced. One of these cDNAs, TALT5j, was used for further study.

To obtain a cDNA that includes the region of EMMPRIN that is 3’ to peptide #51, we used the rapid amplification of cDNA ends protocol as described previously (24). A pool of cDNA was prepared from poly(A)+ RNA of LX-1 cells using the dT17 adapter primer E of 5’-GAATTCGAGATCC-GATACTCTTTTTTTTTTTTTTTTTT. The reaction mixture was then amplified

```
GAGGTGAC-TCC-ATT-GAG-TAC-CTG-GAG-CCG-GAT-GCC-GAG-AAG-GGC-GAG
```

in PCR using the dT17 adapter primer E and primer B, derived from peptide #51 (Table 1). After ligating the PCR products into the pCRII vector, transforming E. coli, and selecting white colonies, the plasmid DNA was isolated, and the inserts were sized and sequenced. One of these cDNAs, TALT3g, was used for further study.

Construction of a cDNA Encoding the Complete EMMPRIN Sequence. The inserts of the 5’ and 3’ cDNA clones for EMMPRIN (TALT5j and TALT3g) overlapped by 20 nucleotides and thus were used in the PCR-based, overlap extension technique (25) to yield a single cDNA with the complete open reading frame for EMMPRIN.

First, the cDNAs corresponding to the 5’ region (TALT5j) and the 3’ region (TALT3g) were amplified by PCR in two separate reactions. For TALT5j, a 30-mer sense primer, composed of a 13-bp BamHI adapter at the 5’ end followed by a 17-mer corresponding to the 5’ terminus of TALT5j, was used (primer F, 5’-GCCCATACCCTCCGAGGATAGATTACG). The anti-sense primer was a 20-mer which corresponded to the 3’ terminus of TALT5j (primer B’ in Table 1). For TALT3g, a 20-mer sense primer corresponding to the 5’ terminus of TALT3g (primer B in Table 1) and the dT17, adapter primer (5’-GAATTCGAGATCC-GATACTCTTTTTTTTTTTTTTTTTTTTTT) were used. The amplified products were electrophoresed in agarose, and the bands corresponding to 0.6 kb (TALT5j) or 1.1 kb (TALT3g) were identified by ethidium bromide staining and excised.

In the third and final reaction, the gel-purified products from both the above reactions were used as templates for fusion by overlap extension (25). For this reaction, the 30-mer primer F (5’-most primer), used above for amplification of TALT5j, was used with the dT17, adapter primer E (3’-most primer). The 1.6-kb DNA product that was generated in this PCR was ethanol precipitated, agarose gel-purified, and subcloned into pBluescript. Sequencing of this cDNA confirmed its identity with the combined sequence of the two separate cDNAs. The complete sequence is shown in Fig. 2.

**Northern Blot Analysis.** Northern blot analysis was performed as described previously (16). Ten μg of total RNA was electrophoresed on a 1% agarose gel and transferred to nitrocellulose membrane. The blot was then hybridized overnight with EMMPRIN cDNA, which had been radiolabeled by nick-translation with 32P-labeled dCTP. After washing, the filter was exposed to Kodak XR-5 film at −80°C for 48 h.

**Deletion Analyses.** The strategy used for making deletion constructs was adapted from that used to create the full-length EMMPRIN cDNA, i.e., overlap extension (25). The primers used for the following reactions were primer F, 5’-GCCCATACCCTCCGAGGATAGATTACG; primer G, 5’-ACG-GAGGCTCCCGGGAGGAGGAGGAATCTG; primer H, 5’-AGCCG-CTTCCGAGAAGGCACCAGTGCGCCGCTTCT; primer I, 5’-GGCGGCGGCGCAGAGGAGGAGGAGGAATCTG; and primer J, 5’-TCTTACCT- ACTAGAAGCCCGGCAAA-3’.

The extracellular immunoglobulin domain I deletion construct was made as follows: (a) two PCRs, using EMMPRIN cDNA as template, were used to synthesize DNA fragments on each side of the desired deletion, one with the primers F (a BamHI adapter plus the 5’ end of the EMMPRIN cDNA sequence) and reverse complement of G, and the other with primer G and the dT17, adapter primer E. Primer G is composed of sequences from each side of the desired deletion, i.e., nucleotides 50–63 linked directly to 319–336 (Fig. 2). PCR products were analyzed and purified on 1% low-melting point agarose. DNA of expected sizes, 108 bp from the first reaction and 1306 bp from the second reaction, were cut from the gels and incubated in 300 μl H2O at
65°C for 10 min; (b) a PCR was performed with 10 ng each of these two fragments, plus primer F and the dT17 adapter. The PCR product was precipitated with 0.3 M ammonium acetate and 2 volumes of ethanol, EcoRI, and electrophoresed on 1% low-melting point agarose. DNA with the expected size, 1382 bp, was cut from the gel and purified on Spin bind extraction cartridge; and (c) the purified DNA was ligated into pBluescript, and used to transform XL-1 blue cells.

Fig. 2. Sequence of human EMMPRIN cDNA. Nucleotide sequence and derived amino acid sequence for human EMMPRIN; nucleotides are numbered on the left and amino acids on the right. The putative signal and transmembrane amino acid sequences and the polyadenylation signal are underlined. The shaded regions indicate the deduced amino acid sequences that match the six peptides derived from EMMPRIN. The 20-nucleotide overlapping region of the 5' and 3' cDNAs, TALT5j and TALT3j, is indicated by a line above nucleotide positions #517–536. The stop codon is marked by an asterisk.

Amino Acid Sequences of Peptides Derived from EMMPRIN. Previously, we have reported the amino acid sequence of four peptides termed #26, #51, #59, and #61, obtained by HPLC fractionation of tryptic peptide digests of EMMPRIN (14). Subsequently, we have sequenced two additional peptides, #38 and #74, from EMMPRIN. The sequences of the six peptides are shown in Table 2.

Isolation of EMMPRIN cDNAs. Several attempts to obtain EMMPRIN cDNA clones by screening an LX-1 library with degenerate or best-guess oligonucleotide probes, derived from the peptide sequences, were unsuccessful. Therefore, the following steps were taken to obtain a cDNA for EMMPRIN. First, a small EMMPRIN-specific cDNA corresponding to a single peptide-derived sequence, peptide #51, was generated by RT-PCR with degenerate primers. Isolation of this cDNA confirmed the presence of an mRNA containing the 3′ end of 2 X SDS-PAGE sample buffer. The cell extracts were separated on 15% SDS-PAGE and then analyzed for cross-reactivity with the monoclonal antibody E11F4 by Western blotting (26).

RESULTS

Table 2 Amino acid sequences of peptides derived from EMMPRIN. Peptides #38 and #74 (marked with an asterisk) were reported here for the first time; peptides #26, #51, #59, and #61 were reported previously (14).
ing the EMMPRIN-derived sequence in LX-1 cells and provided us with a correct cDNA probe to be used for generation of larger EMMPRIN cDNAs.

After sequencing the small cDNA corresponding to peptide #51, a unique reverse complement primer derived from it was used in RT-PCR, together with a degenerate primer made from the amino terminal peptide #59, to obtain a longer cDNA. Sequencing of one of the isolated cDNAs (TALT5j; insert size, 0.6 kb) revealed the following characteristics: (a) as expected, the sequences at the 5' and 3' ends corresponded to the two primers, D and B', which were based on portions of peptides #59 and #51, respectively; and (b) the complete sequences encoding peptides #26, #61, and #74 (Table 2) were present within the cDNA, demonstrating that the cDNA encodes an authentic EMMPRIN sequence.

An unexpected result from the above approach was that although the sequence of primer D was present at the 5' end of the cDNA, the complete sequence corresponding to peptide #59, from which D is derived, was found to begin 72 nucleotide residues downstream from the primer sequence. Thus, it appears that primer D annealed with a region of the 5' untranslated sequence of EMMPRIN and amplified a cDNA that begins in the 5' untranslated region. A possible explanation of this event comes from comparison of the sequence of this cDNA with sequences recently obtained from an EMMPRIN cDNA derived from a human keratinocyte cDNA library. The sequence at the 3' end of primer D (Table 1) corresponds exactly with a sequence within the 5' untranslated region, namely GAGGA, for the keratinocyte-derived as well as LX-1-derived EMMPRIN cDNAs. Therefore, this unexpected circumstance led to additional information about the 5' untranslated end of our cDNA.

Thus, in summary, the 5' cDNA (TALT5j) corresponds to nucleotide residues −14 to 536, defining +1 as A in the ATG start codon (Fig. 2). It begins with a portion of 5' untranslated sequence, then contains a methionyl initiation codon at the start of a sequence encoding a region with the properties of a signal peptide (amino acid residues #1−21 in Fig. 2). This is followed by the sequence encoding peptide #59 (residues #22−36), corresponding to the NH2-terminus of the mature protein. The cDNA continues through to the 3' end primer, B', encoding part of the amino acid sequence corresponding to peptide #51 (residues #173−178).

The rapid amplification of cDNA ends technique (24) was applied to obtain the 3' cDNA, once more using a primer based on part of the authentic nucleotide sequence from the peptide #51 cDNA, together with a universal 3' oligo d(T)17 adapter primer. Sequence analysis of the PCR products indicated that one of the inserts (TALT3g), 1.1 kb in size, begins at the 5'-terminus with the sequence corresponding to primer B, continues with the sequence of the COOH-terminus of the mature protein. The cDNA continues through to the 3' end primer, B', encoding part of the amino acid sequence corresponding to peptide #51 (residues #173−178).

The rapid amplification of cDNA ends technique (24) was applied to obtain the 3' cDNA, once more using a primer based on part of the authentic nucleotide sequence from the peptide #51 cDNA, together with a universal 3' oligo d(T)17 adapter primer. Sequence analysis of the PCR products indicated that one of the inserts (TALT3g), 1.1 kb in size, begins at the 5'-terminus with the sequence corresponding to primer B, continues with the sequence of the COOH-terminus of the mature protein. The cDNA continues through to the 3' end primer, B', encoding part of the amino acid sequence corresponding to peptide #51 (residues #173−178).

The rapid amplification of cDNA ends technique (24) was applied to obtain the 3' cDNA, once more using a primer based on part of the authentic nucleotide sequence from the peptide #51 cDNA, together with a universal 3' oligo d(T)17 adapter primer. Sequence analysis of the PCR products indicated that one of the inserts (TALT3g), 1.1 kb in size, begins at the 5'-terminus with the sequence corresponding to primer B, continues with the sequence of the COOH-terminus of the mature protein. The cDNA continues through to the 3' end primer, B', encoding part of the amino acid sequence corresponding to peptide #51 (residues #173−178).

The rapid amplification of cDNA ends technique (24) was applied to obtain the 3' cDNA, once more using a primer based on part of the authentic nucleotide sequence from the peptide #51 cDNA, together with a universal 3' oligo d(T)17 adapter primer. Sequence analysis of the PCR products indicated that one of the inserts (TALT3g), 1.1 kb in size, begins at the 5'-terminus with the sequence corresponding to primer B, continues with the sequence of the COOH-terminus of the mature protein. The cDNA continues through to the 3' end primer, B', encoding part of the amino acid sequence corresponding to peptide #51 (residues #173−178).

**Analysis of EMMPRIN cDNA Sequences.** The Northern blots indicate that the EMMPRIN composite cDNA corresponds to all but about 100 nucleotides of the mRNA. Since the clone contains a poly(A) tail, this means that the 100 bases are probably located at the 3'-end of the 5' untranslated region, making the total untranslated region about 115 nucleotides in length.

The CDNA encodes a 269-amino acid residue polypeptide that contains a putative signal peptide of 21 amino acid residues, an extracellular domain of 185 amino acid residues, a putative transmembrane region (residues 206−229), and a carboxy-terminal cytoplasmic domain of 39 amino acid residues (230−269). The transmembrane region includes three leucines (residues 206, 213, and 220) and a phenylalanine (residue 227), occurring every seventh residue, a characteristic feature of the leucine zipper motif (Fig. 2). The extracellular region contains four cysteinyl residues spaced in a manner that gives rise to two distinct domains with the characteristics of proteins in the immunoglobulin superfamily. These residues in EMMPRIN are located at amino acid residue positions 41, 87, 126, and 185.

**Expression of Recombinant EMMPRIN Protein and Deletion Constructs.** Although the EMMPRIN cDNA encodes amino acid sequences identical to those of peptides directly isolated from immunoadfinity-purified EMMPRIN protein, we reconfirmed the identity of the EMMPRIN cDNA by showing that the recombinant protein is recognized by activity-blocking monoclonal antibody. To accomplish this, the recombinant protein was expressed in pBluescript and then assayed by Western blotting with E11F4, a monoclonal antibody that blocks the activity of EMMPRIN protein from LX-1 cells (13, 14). As shown in Fig. 4B, Lane 2, E11F4 reacts with the recombinant EMMPRIN protein. Two immunoreactive bands were obtained; these correspond in size to the two forms of EMMPRIN noted previously (14) and most likely arise by proteolysis.

The bacterial recombinant protein was also used to determine the approximate location of the epitope for E11F4, taking advantage of the lack of posttranslational modifications that would interfere with such studies on the native, immunoadfinity-purified protein. Modified EMMPRIN expression plasmids were made containing deletions in four locations. As seen in Fig. 4A, these deletions were: (a) deletion of the extracellular immunoglobulin domain I (ΔECI); (b) deletion of the extracellular immunoglobulin domain II (ΔECII); (c) deletion of the transmembrane domain (ΔTM); and (d) deletion of the cytoplasmic domain (ΔCYT). XL-1 blue cells were transformed with the deletion expression pBluescript plasmids, and the expressed proteins were analyzed by SDS-PAGE and Western blotting with the monoclonal antibody E11F4. As seen in Fig. 4B, all the plasmids produce protein that is immunoreactive, except for the plasmid lacking immunoglobulin domain I.
Fig. 4. Localization of EIIF4 epitope in recombinant EMMPRIN. A, strategy for deleting the four major domains of the EMMPRIN cDNA. ΔECI and ΔECII, deletion of one of two immunoglobulin-like domains within the extracellular region; ΔTM, deletion of transmembrane domain; ΔCYT, deletion of cytoplasmic domain; SP, signal peptide; B, Western blot. The transformed XL-1 blue cells were incubated in the absence (Lanes 1, 3, 5, 7, and 9) or presence (Lanes 2, 4, 6, 8, and 10) of 0.5 mM isopropyl-β-D-thiogalactopyranoside and processed for Western blotting. Lanes 1 and 2, recombinant EMMPRIN without deletion; Lanes 3 and 4, ΔECI; Lanes 5 and 6, ΔECII; Lanes 7 and 8, ΔTM; Lanes 9 and 10, ΔCYT. Molecular weight marker positions in kilodaltons are shown.

These results demonstrate that our cDNA encodes the protein which is reactive with our activity-blocking monoclonal antibody and that the antibody epitope exists in the extracellular immunoglobulin domain I. This, in turn, implies that the functional site of the metalloproteinase stimulatory activity of EMMPRIN is likely to be localized to sequences contained in the immunoglobulin domain I region.

DISCUSSION

We have isolated and fused two overlapping cDNA clones, using the polymerase chain reaction, which together encode the complete, 269-amino acid open reading frame for EMMPRIN. The identity of the clones was confirmed by comparison to several peptide sequences derived from immunoaffinity-purified EMMPRIN. Recognition of the translation product by the activity-blocking monoclonal antibody EIIF4 further confirms the identity of the cDNAs as the desired EMMPRIN clones. In addition, we have recently isolated an EMMPRIN cDNA from a human keratinocyte cDNA library. The open reading frame of the latter cDNA has an identical sequence to the cDNAs as the desired EMMPRIN clones. In addition, we have recently isolated an EMMPRIN cDNA from a human keratinocyte cDNA library. The open reading frame of the latter cDNA has an identical sequence to the cDNA obtained in the present study by PCR-based techniques, except for two nucleotide residues; however, the deduced amino acid sequences are identical for the two cDNAs.7 Finally, we have recently demonstrated that recombinant EMMPRIN isolated from CHO cells transfected with EMMPRIN cDNA stimulates metalloproteinase production in fibroblasts.9

The composite cDNA obtained in the present study has a small 5'-untranslated region, followed by an initiation codon and sequences that have the properties of a signal peptide sequence, when using the rules of von Heijne (27). The subsequent codons agree perfectly with our amino terminal peptide sequence for the mature protein, demonstrating that the signal peptide sequence is genuine. The 248 codons after the signal sequence encode a 185-amino acid extracellular domain consisting of two regions characteristic of the immunoglobulin superfamily, followed by 24-amino acid residues comprising the transmembrane domain and a 39-amino acid cytoplasmic domain. The 248-amino acid residues of the mature protein correspond to an approximate molecular weight of 27,000. However, the purified protein has a larger molecular weight of ~58,000 (13). This difference is mainly due to glycosylation of the protein7 (20).

We have shown previously that EMMPRIN is present at the surface of tumor cells (13) and has the properties of a membrane-intercalated protein (12). On the basis of these findings, we proposed previously that EMMPRIN is attached to the plasma membrane via a transmembrane domain and interacts with a receptor on fibroblasts via an extracellular domain (1). The presence in the cDNA of sequences typical of a signal peptide and a transmembrane region is consistent with EMMPRIN being an integral plasma membrane protein. After the termination codon, the cDNA contains a 3' untranslated region ending in a poly(A) tail. Northern blot analysis indicates that the mRNA for EMMPRIN is ~1.7 kb in size, which is approximately the same as that of the fused EMMPRIN cDNA. It is evident, however, that a portion of the 5' untranslated region is lacking from this cDNA.

The EMMPRIN cDNA sequences were used in computer searches of the EMBL and GenBank data bases to detect homology with other known proteins. These searches revealed that the EMMPRIN cDNA is identical to two other human cDNAs, encoding proteins of un-
known function, basigin (19) and the M6 antigen (20). Therefore, our studies, which have been performed from a functional standpoint over the course of many years (4, 6, 12–16), have elucidated at least one biological function of these molecules. We are changing our former designation of the molecule from TCSF to EMMPRIN and would suggest that all the above proteins now be designated EMMPRIN, because the acronym more accurately implies at least one definitive function of the glycoprotein, i.e., stimulation of MMP synthesis via cell-cell interaction.

The fact that EMMPRIN is a member of the immunoglobulin superfamily is also compatible with the idea that, in similar fashion to the N-CAM, I-CAM, and other related subgroups of the immunoglobulin superfamily (28), it acts via cell-cell interactions (1). We are currently attempting to identify the molecule on the surface of fibroblasts that interacts with tumor cell-derived EMMPRIN, causing increased fibroblast MMP production. Our recent finding that EMMPRIN is expressed in keratinocytes and localized in the basal layers of the epidermis suggests the possibility that EMMPRIN may have a natural function in embryonic development or wound healing by causing dermal fibroblasts to increase their MMP production, thus facilitating tissue remodeling (18). The antibody to M6 antigen localizes EMMPRIN to granulocytes in patients with rheumatoid arthritis (20), possibly indicating a role for EMMPRIN in stromal MMP production and the consequent matrix degradation that occurs in the arthritic joint. Thus, we propose that EMMPRIN and related molecules are important mediators of matrix remodeling in normal and pathological tissues.

With respect to tumorigenesis, it has become clear that: (a) MMPs are crucial to the process of tumor cell invasion through basement membranes and interstitial matrices (1–3); and (b) in the case of interstitial collagenase, stromelysin, and M, 72,000 gelatinase (type IV collagenase), the MMPs involved are produced mainly by peritumoral fibroblasts rather than by the tumor cells themselves (8–11). Since tumor cell-derived EMMPRIN causes a significant increase in the levels of these three enzymes in human fibroblasts (6, 13–16) and since EMMPRIN is associated with the surface of many types of tumor cells in vivo and in vitro (13, 17), it is very likely that EMMPRIN is a central factor in the stimulation of MMPs required for tumor invasion and metastasis.

ACKNOWLEDGMENTS

We thank Drs. Marion Gordon and Bryan Toole for their helpful suggestions and encouragement during various parts of the project and preparation of the manuscript; Dr. Nicholas Grammatikakis for suggesting the name EMMPRIN; Dr. William S. Lane for preparing primer E; and Diane H. Silva for her technical help and assistance during preparation of the manuscript.

REFERENCES

The Human Tumor Cell-derived Collagenase Stimulatory Factor (Renamed EMMPRIN) Is a Member of the Immunoglobulin Superfamily

Chitra Biswas, Ying Zhang, Rosana DeCastro, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/55/2/434

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/55/2/434. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.