Primary structure of matrilin-3, a new member of a family of extracellular matrix proteins related to cartilage matrix protein (matrilin-1) and von Willebrand factor

Raimund Wagener*, Birgit Kobbe, Mats Paulsson

Institute for Biochemistry II, Medical Faculty, University of Cologne, Joseph-Stelzmann-Strasse 52, D-50931 Cologne, Germany

Received 4 June 1997; revised version received 2 July 1997

Abstract A mouse cDNA encoding for matrilin-3, the third member of the novel matrilin family of extracellular matrix proteins, was cloned. The protein precursor of 481 amino acids consists of a putative signal peptide, a short positively charged sequence, a single vWFA-like domain followed by four epidermal growth factor-like modules and a potential coiled-coil α-helical oligomerization domain at the C-terminus. It is the smallest member of the matrilin family with a predicted M_r of the mature protein of 48 902. The primary structure of a C-terminal portion of 310 amino acids of the human matrilin-3 was determined and showed a sequence identity to the mouse matrilin-3 of 84.8%. Northern blot hybridization of mouse matrilin-3 mRNA showed a 2.9 kb mRNA expressed in sternum, femur and trachea and indicates a cartilage-specific expression.

Key words: Cartilage matrix protein; Matrilin; Extracellular matrix; Von Willebrand factor type A; Epidermal growth factor; Coiled-coil α-helix

1. Introduction

The multidomain structure is a common feature of many extracellular proteins [1] and approximately 50 different domains have been identified in various animal proteins [2]. A unique protein is formed by the tandem arrangement of a certain set of domains. The protein superfamily containing the von Willebrand factor type A (vWFA)-like domain consists of proteins with different domain organization and a variety of functions [3]. The vWFA-like domain was first described in von Willebrand factor where it plays a key role in promoting platelet adhesion to the subendothelium. Several vWFA-like domains have been implicated in interactions with collagen [3].

Recently a new subfamily of extracellular matrix proteins with vWFA-like domains was defined — the matrilins. To date it consists of two members, cartilage matrix protein (CMP) now alternatively named matrilin-1 and the recently cloned matrilin-2 [4]. Matrilin-1 is a trimer of identical ellipsoid subunits assembled via their C-terminal extension domains in a coiled-coil α-helix [5]. Aggrecan [6,7] and type II collagen-containing fibrils [8] have been proposed as binding partners for matrilin-1 and it has been suggested that matrilin-1 might play an integrating role in cartilage extracellular matrix organisation as a bridging molecule between these two major constituents [7]. It has also been shown that matrilin-1 can form filamentous networks independent of collagen fibrils [9]. Matrilin-1 is expressed only in some types of the hyaline cartilage [10] whereas matrilin-2 is found in calvaria, uterus, heart, skeletal muscle, brain and skin but not in normal cartilage [4]. The expression patterns appear to be complementary. The primary structure of matrilin-1 has been deduced from chicken [11], man [12] and mouse [13] cDNA and that of matrilin-2 from mouse and partially from human cDNA [4]. In both cases the sequence is highly conserved between these species. The modular structure of the two is very similar with each containing two vWFA-like domains, interrupted by 1 or 10 EGF-like domains, respectively, and a C-terminal oligomerization domain. Matrilin-2 possesses an additional unique sequence of 75 amino acids between the second vWFA-like domain and the oligomerization unit and has a highly positively charged stretch of 16 amino acids containing 6 arginine residues between the signal peptide and the first vWFA-like domain. It has been proposed from the structural homologies and the complementary sites of expression that these two matrilins may have a similar function in different forms of extracellular matrix [4].

Supposing the existence of more members of this new family of extracellular matrix proteins, we looked for gene products with a similar modular structure by searching in the EST databases. Here we report on the deduced primary structure of matrilin-3 from mouse, which is highly homologous to matrilin-1 and -2 and differs from both by lacking of the second vWFA-like domain.

2. Materials and methods

2.1. Clones and libraries

The matrilin homologue human lung EST clone, IMAGE Consortium Clone ID 119728 [14], was obtained from ATCC (#341513). An oligo(dT)- and random-primed lung cDNA library in lambda ZAP® II vector from 6–8-week-old female mice (B6CBA) was purchased from Stratagene.

2.2. Screening of library and DNA sequencing

2×10^6 pfu were screened using a 1.1 kb N-terminal fragment of the human EST clone. The hybridization conditions were 5×SSPE (1×SSPE is 0.15 M NaCl, 10 mM sodium phosphate (pH 7.7) and 1 mM EDTA), 5×Denhardt's solution (50×Denhardt's solution is 1% BSA, 1% Ficoll 400 and 1% polyvinylpyrrolidone), 0.5% SDS, 45% formamide and 20 μg/ml salmon sperm DNA at 42°C. Filters were washed in the last washing step with 0.1×SSPE, 0.1% SDS for 15 min at 55°C. The positive clone was in vivo excised yielding a cDNA clone in pBluescript SK(−). Using the dideoxy method and

*Corresponding author. Fax: (49) 221-4786977.
E-mail: akc31@rs1.rrz.uni-koeln.de

Abbreviations: vWFA, von Willebrand factor type A; CMP, cartilage matrix protein; EGF, epidermal growth factor; EST, expressed sequence tag; GAPDH, glyceraldehyde-3-phosphate dehydrogenase

The nucleotide sequences reported in this paper have been submitted to the EMBL nucleotide sequence database with accession numbers Y10521 and Y13341.

PII S0014-5793(97)00895-8

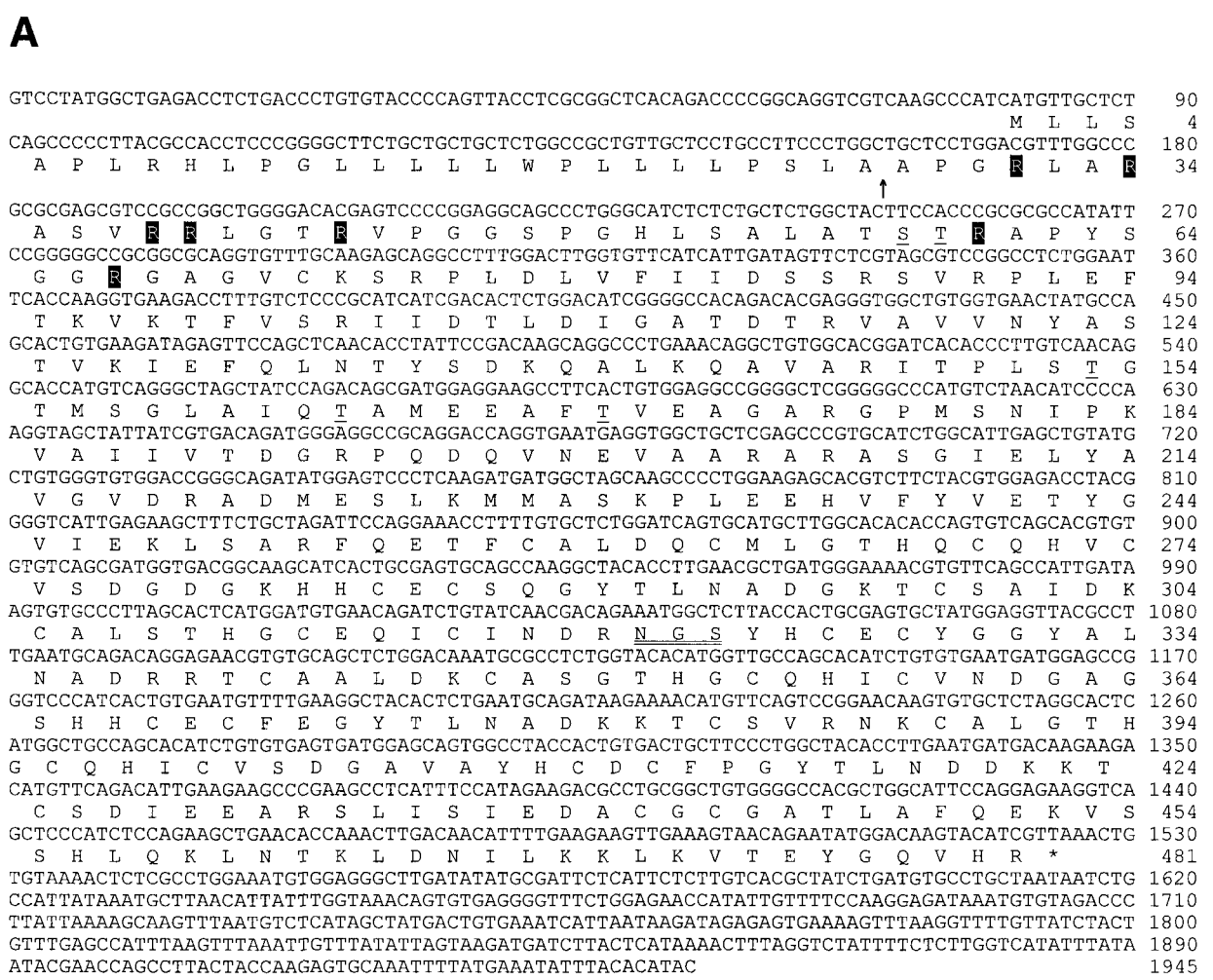

A
GTCCTATGGCTGAGACCTCTGACCCTGTGTACCCCAGTTACCTCGCGGCTCACAGACCCCGGCAGGTCGTCAAGCCCATCATGTTGCTCT 90
M L L S 4
CAGCCCCCTTACGCCACCTCCCGGGGCTTCTGCTGCTGCTCTGGCCGCTGTTGCTCCTGCCTTCCCTGGCTGCTCCTGGACGTTTGGCCC 180
A P L R H L P G L L L L L W P L L L L P S L A A P G R L A R 34
GCGCGAGCGTCCGCCGGCTGGGGACACGAGTCCCCGGAGGCAGCCCTGGGCATCTCTCTGCTCTGGCTACTTCCACCCGCGCGCCATATT 270
A S V R R L G T R V P G G S P G H L S A L A T S T R A P Y S 64
CCGGGGGCCGCGGCGCAGGTGTTTGCAAGAGCAGGCCTTTGGACTTGGTGTTCATCATTGATAGTTCTCGTAGCGTCCGGCCTCTGGAAT 360
G G R G A G V C K S R P L D L V F I I D S S R S V R P L E F 94
TCACCAAGGTGAAGACCTTTGTCTCCCGCATCATCGACACTCTGGACATCGGGGCCACAGACACGAGGGTGGCTGTGGTGAACTATGCCA 450
T K V K T F V S R I I D T L D I G A T D T R V A V V N Y A S 124
GCACTGTGAAGATAGAGTTCCAGCTCAACACCTATTCCGACAAGCAGGCCCTGAAACAGGCTGTGGCACGGATCACACCCTTGTCAACAG 540
T V K I E F Q L N T Y S D K Q A L K Q A V A R I T P L S T G 154
GCACCATGTCAGGGCTAGCTATCCAGACAGCGATGGAGGAAGCCTTCACTGTGGAGGCCGGGGCTCGGGGGCCCATGTCTAACATCCCCA 630
T M S G L A I Q T A M E E A F T V E A G A R G P M S N I P K 184
AGGTAGCTATTATCGTGACAGATGGGAGGCCGCAGGACCAGGTGAATGAGGTGGCTGCTCGAGCCCGTGCATCTGGCATTGAGCTGTATG 720
V A I I V T D G R P Q D Q V N E V A A R A R A S G I E L Y A 214
CTGTGGGTGTGGACCGGGCAGATATGGAGTCCCTCAAGATGATGGCTAGCAAGCCCCTGGAAGAGCACGTCTTCTACGTGGAGACCTACG 810
V G V D R A D M E S L K M M A S K P L E E H V F Y V E T Y G 244
GGGTCATTGAGAAGCTTTCTGCTAGATTCCAGGAAACCTTTTGTGCTCTGGATCAGTGCATGCTTGGCACACACCAGTGTCAGCACGTGT 900
V I E K L S A R F Q E T F C A L D Q C M L G T H Q C Q H V C 274
GTGTCAGCGATGGTGACGGCAAGCATCACTGCGAGTGCAGCCAAGGCTACACCTTGAACGCTGATGGGAAAACGTGTTCAGCCATTGATA 990
V S D G D G K H H C E C S Q G Y T L N A D G K T C S A I D K 304
AGTGTGCCCTTAGCACTCATGGATGTGAACAGATCTGTATCAACGACAGAAATGGCTCTTACCACTGCGAGTGCTATGGAGGTTACGCCT 1080
C A L S T H G C E Q I C I N D R N G S Y H C E C Y G G Y A L 334
TGAATGCAGACAGGAGAACGTGTGCAGCTCTGGACAAATGCGCCTCTGGTACACATGGTTGCCAGCACATCTGTGTGAATGATGGAGCCG 1170
N A D R R T C A A L D K C A S G T H G C Q H I C V N D G A G 364
GGTCCCATCACTGTGAATGTTTTGAAGGCTACACTCTGAATGCAGATAAGAAAACATGTTCAGTCCGGAACAAGTGTGCTCTAGGCACTC 1260
S H H C E C F E G Y T L N A D K K T C S V R N K C A L G T H 394
ATGGCTGCCAGCACATCTGTGTGAGTGATGGAGCAGTGGCCTACCACTGTGACTGCTTCCCTGGCTACACCTTGAATGATGACAAGAAGA 1350
G C Q H I C V S D G A V A Y H C D C F P G Y T L N D D K K T 424
CATGTTCAGACATTGAAGAAGCCCGAAGCCTCATTTCCATAGAAGACGCCTGCGGCTGTGGGGCCACGCTGGCATTCCAGGAGAAGGTCA 1440
C S D I E E A R S L I S I E D A C G C G A T L A F Q E K V S 454
GCTCCCATCTCCAGAAGCTGAACACCAAACTTGACAACATTTTGAAGAAGTTGAAAGTAACAGAATATGGACAAGTACATCGTTAAACTG 1530
S H L Q K L N T K L D N I L K K L K V T E Y G Q V H R * 481
TGTAAAACTCTCGCCTGGAAATGTGGAGGGCTTGATATATGCGATTCTCATTCTCTTGTCACGCTATCTGATGTGCCTGCTAATAATCTG 1620
CCATTATAAATGCTTAACATTATTTGGTAAACAGTGTGAGGGGTTTCTGGAGAACCATATTGTTTTCCAAGGAGATAAATGTGTAGACCC 1710
TTATTAAAAGCAAGTTTAATGTCTCATAGCTATGACTGTGAAATCATTAATAAGATAGAGAGTGAAAAGTTTAAGGTTTTGTTATCTACT 1800
GTTTGAGCCATTTAAGTTTAAATTGTTTATATTAGTAAGATGATCTTACTCATAAAACTTTAGGTCTATTTTCTCTTGGTCATATTTATA 1890
ATACGAACCAGCCTTACTACCAAGAGTGCAAATTTTATGAAATATTTACACATAC 1945
B
GGAGGCAGGGGCTCGAGAGCCCTCTTCTAACATCCCTAAGGTGGCCATCATTGTTACAGATGGGAGGCCCCAGGACCAGGTGAATGAGGT 90
E A G A R E P S S N I P K V A I I V T D G R P Q D Q V N E V 30
GGCGGCTCGGGCCCAAGCATCTGGTATTGAGCTCTATGCTGTGGGCGTGGACCGGGCAGACATGGCGTCCCTCAAGATGATGGCCAGTGA 180
A A R A Q A S G I E L Y A V G V D R A D M A S L K M M A S E 60
GCCCCTAGAGGAGCATGTTTTCTACGTGGAGACCTATGGGGTCATTGAGAAACTTTCCTCTAGATTCCAGGAAACCTTCTGTGCGCTGGA 270
P L E E H V F Y V E T Y G V I E K L S S R F Q E T F C A L D 90
CCCCTGTGTGCTTGGAACACACCAGTGCCAGCACGTCTGCATCAGTGATGGGGAAGGCAAGCACCACTGTGAGTGTAGCCAAGGATACAC 360
P C V L G T H Q C Q H V C I S D G E G K H H C E C S Q G Y T 120
CTTGAATGCCGACAAGAAAACGTGTTCAGCTCTTGATAGGTGTGCTCTTAACACCCACGGATGTGAGCACATCTGTGTGAATGACAGAAG 450
L N A D K K T C S A L D R C A L N T H G C E H I C V N D R S 150
TGGCTCTTATCATTGTGAGTGCTATGAAGGTTATACCTTGAATGAAGACAGGAAAACTTGTTCAGCTCAAGATAAATGTGCTTTGGGTAC 540
G S Y H C E C Y E G Y T L N E D R K T C S A Q D K C A L G T 180
CCATGGGTGTCAGCACATTTGTGTGAATGACAGAACAGGGTCCCATCATTGTGAATGCTATGAGGGCTACACTCTGAATGCAGATAAAAA 630
H G C Q H I C V N D R T G S H H C E C Y E G Y T L N A D K K 210
AACATGTTCAGTCCGTGACAAGTGTGCCCTAGGCTCTCATGGTTGCCAGCACATTTGTGTGAGTGATGGGGCCGCATCCTACCACTGTGA 720
T C S V R D K C A L G S H G C Q H I C V S D G A A S Y H C D 240
TTGCTATCCTGGCTACACCTTAAATGAGGACAAGAAAACATGTTCAGCCACTGAGGAAGCACGAAGACTTGTTTCCACTGAAGATGCTTG 810
C Y P G Y T L N E D K K T C S A T E E A R R L V S T E D A C 270
TGGATGTGAAGCTACACTGGCATTCCAGGACAAGGTCAGCTCGTATCTTCAAAGACTGAACACTAAACTTGATGACATTTTGGAGAAGTT 900
G C E A T L A F Q D K V S S Y L Q R L N T K L D D I L E K L 300
GAAAATAAATGAATATGGACAAATACATCGTTAAATTGCTCCAATTTCTCACCTGAAAATGTGGACAGCTTGGTGTACTTAATACTCATG 990
K I N E Y G Q I H R * 310
CATTCTTTTGCACACCTGTTATTGCCAATGTTCCTGCTAATAATTTGCCATTATCTGTATTAATGCTTGAATATTACTGGATAAATTGTA 1080
TGAAGATCTTCTGCAGAATCAGCATGATTCTTCCAAGGAAATACATATGCAGATACTTATTAAGAGCAAACTTTAGTGTCTCTAAGTTAT 1170
GACTGTGAAATGATTGGTAGGAAATAGAATGAAAAGTTTAGTGTTTCTTTATCTACTAATTGAGCCATTTAATTTTTAAATGTTTATATT 1260
AGATAACCATATTCACAATGGAAACTTTAGGTCTAGTTTCTTTTGATAGTATTTATAATATAAATCAATCTTATTACTGAGAGTGCAAAT 1350
TGTACAAGGTATTTACACATAC 1372

Fig. 1. Nucleotide and deduced amino acid sequences of matrilin-3. A: Complete sequence of mouse matrilin-3 precursor. Arrowhead: Predicted propeptidase cleavage site. Arginine residues at the amino terminus of the protein are shaded black. The predicted *O*-glycosylation sites (amino acid residues 58, 59, 153, 163 and 170) are underlined once, the potential N-linked glycosylation site (amino acid residues 321–323) is underlined twice. B: Partial nucleotide and amino acid sequence of human matrilin-3.
←

an ALF automatic sequencer (Pharmacia) the plasmids were sequenced in both directions with universal and internal primers. Nucleotide sequence analysis was performed with the programs of the GCG package [15].

2.3. Northern blot analysis

Total RNA of various tissues was extracted from 4-week-old female mice by the guanidinium-thiocyanate method. Poly(A^+) RNA was prepared from the total RNA using the Pharmacia mRNA kit. Aliquots (3–6 μg) of poly(A^+) RNA were electrophoresed on 1.2% denaturing agarose gel, blotted and hybridized. The conditions for the last washing step were: 0.1×SSPE, 0.1% SDS at 65°C for 15 min.

3. Results

A search in the EST database yielded an EST from human lung (T94707) with a sequence identity to the C-terminal part of the N-terminal vWFA-like domains of matrilin-1 and -2 of 59.8% and 48.3%, respectively. We sequenced the EST clone and found an open reading frame (ORF) of 310 amino acids (Fig. 1B). The clone was incomplete, starting within a vWFA-like domain followed by four EGF-like domains and a putative coiled-coil α-helical domain (Fig. 2). From the striking homologies to the other matrilins we concluded that we had found a new member of the matrilin family. From this point onwards the work was continued in the mouse. To get the complete sequence we screened a mouse lung cDNA library with the human clone as a probe. We isolated a clone that contained an ORF of 1443 bp ending with a TAA-stop codon preceded by a 80 bp long 5′-untranslated region (Fig. 1A) with a sequence identity to the human EST of 84.8% on the protein level indicating that it codes for the mouse homologue. The nucleotide sequence codes for a protein precursor of 481 amino acids with a putative signal peptide of 27 amino acids as predicted by a method using neural networks [16]. An in frame ATG codon is found further upstream which, if used, would yield a 25 residue longer signal peptide. However, this would result in an unusual signal peptide sequence followed by a classical signal peptide. The mature secreted protein has a predicted M_r of 48902 and is to date the least complex member of the matrilin family (Fig. 2). It contains only one vWFA-like domain followed by four EGF-like domains and a putative oligomerization domain. N-terminal to the vWFA-like domain is a positively charged sequence of about 40 amino acid residues with a high content of arginine. The second EGF-like domain of the mouse matrilin-3 contains one potential N-glycosylation site (Asn^{321}–Gly–Ser) which is, however, lacking in the partial human sequence. By using a neural networks prediction method [17] we found two potential *O*-glycosylation sites in the positively charged sequence and three in the vWFA-like module of mouse matrilin-3 (Fig. 1A). The corresponding sequence of the human matrilin-3 has not been determined.

A sequence alignment of the different modules with their counterparts in the other matrilins shows the striking homology (Fig. 3). The sequence identity to the first vWFA-like domains of mouse matrilin-1 and -2 is 58.5% and 49.7%, respectively, and for the second vWFA-like domain 37.8% and 41.2%. The metal ion-dependent adhesion site (MIDAS) motif [18], the 2 flanking cysteine residues and the 6 hydrophobic residues that are highly conserved in vWFA-like domains [19] are also conserved in matrilin-3. The structure consists of alternating amphipatic α-helices and hydrophobic β-strands (Fig. 3A) and is in good agreement with the structure of the vWFA-like domain of the α-subunit of the A domain of the integrin CR3 which has been determined at high resolution [18]. The EGF-like domains, which are characterized by the spacing of 6 cysteines, have an average sequence identity of 50.3% to the EGF-like domain of matrilin-1 and range between 55% for the third and 47.5% for the second EGF-like domain. Comparing the EGF-like domains of matrilin-2 and -3, the identity is highest (61.5%) between the third EGF-like domain of matrilin-3 and the third EGF-like domain of matrilin-2 and lowest (36.6%) between the fourth EGF-like domain of matrilin-3 and seventh EGF-like domain of matrilin-2. The average sequence identity is 46%. The EGF-like domains of matrilin-3 all have an additional asparagine 3 residues C-terminal of the third cysteine residue and lack the key residues involved in Ca^{2+} binding to EGF-like domains (Fig. 3B). The potential oligomerization domain has the lowest homology to the other matrilins (Fig. 3C). While the degree of identity is 38.1% to matrilin-1 and only 21.7% to matrilin-2 the positions 'a' and 'd' of the heptad repeats are well conserved. The COILS program [20] detects one heptad repeat less than in the other matrilins. A threonine residue instead of a hydrophobic amino acid is found at the 'd' position at the beginning and at the end of the coiled-coil domain. The pair of closely spaced cysteine residues at the beginning of the oligomerization domain is conserved in all three matrilins. In the case of matrilin-1 it was shown that these cysteines form interchain disulfide bonds [21].

We studied the expression of the *matrilin-3* gene in different mouse tissues by Northern hybridization with poly(A^+) RNA. A 2.9 kb band was detected in sternum, femur and very weakly also in trachea. Other tissues tested were negative (Fig. 4).

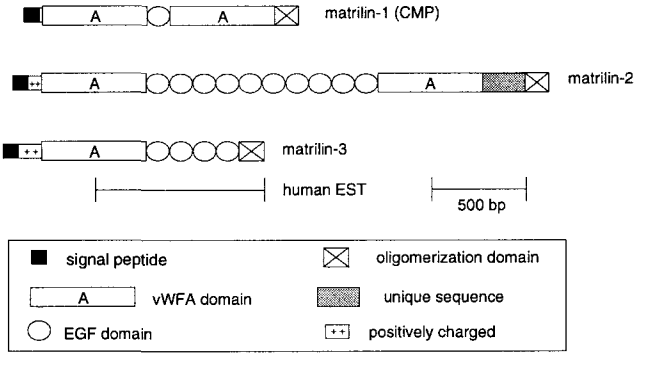
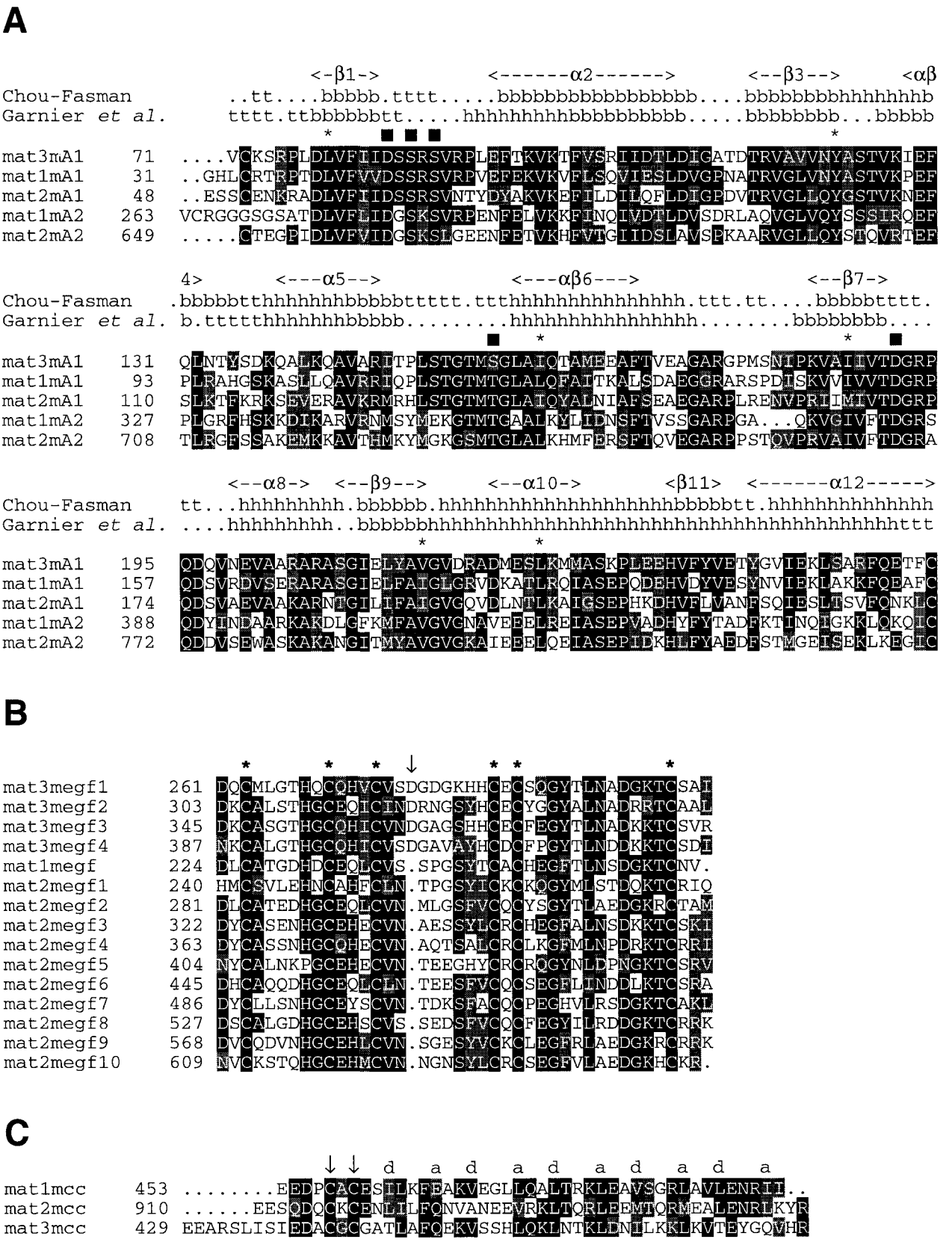


Fig. 2. Comparison of the modular structure of the matrilin family. The horizontal line denotes the coding region of the incomplete human matrilin-3 cDNA clone (ID 119728) [14].



4. Discussion

We have isolated and characterized a cDNA for matrilin-3, the third member of a novel family of extracellular matrix proteins, as evidenced from the striking sequence similarity of the protein to the other two members which are the nearest homologues detected in the databases. Matrilin-3 shares the modular composition of vWFA-like, EGF-like and coiled-coil

Fig. 3. Amino acid sequence alignments of the matrilin modules. Murine sequences were aligned by the PILEUP program of the GCG package, using the default parameters. Matrilin-1 (CMP) [13] is numbered from the first codon identified, matrilin-2 [4] and -3 are numbered from the first amino acid of the protein precursor. A: Sequence alignment of the vWFA-like modules. The locations of the α-helices and β-sheets determined from averaged structure predictions of 75 modules [19] are indicated by the arrow ranges β1–α12. Structure predictions by the Chou-Fasman and Garnier methods for the vWFA-like module of matrilin-3 are shown underneath (b, β-sheet; a, α-helix; t, turn). The conserved metal ion-dependent adhesion site [18] and the conserved hydrophobic moieties [19] are denoted with (■) or (*) respectively. B: Sequence alignment of the EGF-like modules. The conserved positions of the cysteine residues are marked by asterisks; the additional aspartic acid residue of the matrilin-3 EGF-like repeats is marked by an arrow. C: Sequence alignment of the α-helical coiled-coil domains. The positions *a* and *d* of the heptad repeats are indicated above. The conserved cysteine residues at the N-terminal end of the coiled-coil region are marked with arrows. ←

domains. Although there are strong homologies between matrilin-3 and the previously known matrilins there are clear differences in the domain arrangement. The expression of the gene is restricted to cartilage and appears similar to the expression pattern of matrilin-1 (CMP).

A comparison of the EGF-like repeats of matrilin-3 with those of matrilin-1 and -2 also shows a remarkable difference in the presence of an additional aspartic acid residue in each EGF-like domain of matrilin-3 (Fig. 3C). From the 3-dimensional structure of a pair of EGF-like domains of fibrillin [22] it can be inferred that the aspartic acid residue would be positioned in the second loop with its side chain exposed to the solvent potentially allowing ligand interactions. The conservation of the MIDAS motif [18] in the vWFA-like domain implies the dependence on divalent cations for function. The presence of heptad repeats indicates the oligomerization of subunits via a coiled-coil α-helical structure. The number of subunits in the oligomer cannot be predicted from the sequence alone. It was recently demonstrated that the exchange of a single amino acid residue in the coiled-coil domain of matrilin-1 leads to the assembly of a tetramer instead of a trimer [23].

The differences between the domain compositions of matrilin-1, -2 and -3 must have implications for their functions.

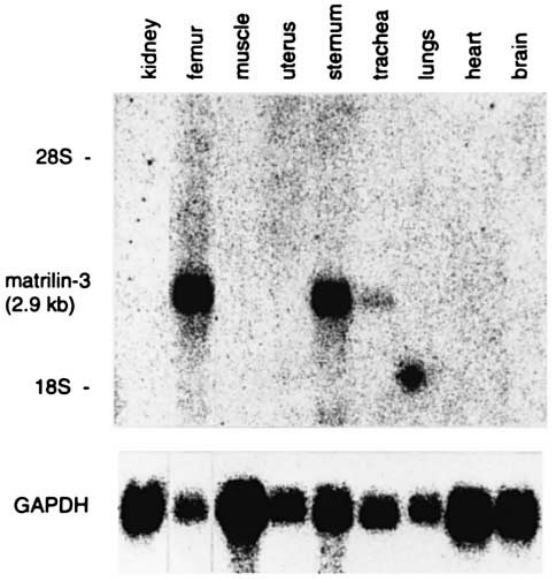


Fig. 4. Distribution of matrilin-3 mRNA in various mouse tissues. Northern hybridization of 3–6 μg poly(A^+) RNA from 4-week-old mice. The blot was hybridized with a 500 bp fragment of the coding region of matrilin-3 cDNA (upper part) and subsequently with a GAPDH cDNA fragment (lower part) to estimate the relative abundance of the matrilin-3 message. The signal in the lane for lungs is due to an artifact. The exposure times for the autoradiography were 12 h using a phosphor imager (Molecular Dynamics).

There is evidence for the binding of matrilin-1 not only to aggrecan [6,7] but also to collagen type II [8]. By use of truncated forms of recombinant matrilin-1 (CMP) it was shown that a binding site involved in assembly into both the collagenous and the non-collagenous fibrils is located in the NH_2-terminal portion of the protein [9]. Matrilin-2 has not yet been studied in this respect. The functions or the potential binding partners of the second vWFA-like domains in matrilin-1 and -2 are not known and it is therefore not possible to make suggestions by analogy about the consequences of the lack of this domain in matrilin-3. Based on the extent of sequence identity it could be that the sites for matrix binding are located on the single vWFA-like domain of matrilin-3 and that a functional difference between matrilin-3 and the matrilins-1 and -2 depends on the unknown function of the lacking vWFA-like domain.

The apparently cartilage-specific expression of matrilin-3 mRNA is analogous to that of matrilin-1. Matrilin-3 mRNA is not detected in matrilin-2 expressing tissues. Although both matrilin-3 cDNA clones described here were derived from lung cDNA libraries we could not detect matrilin-3 RNA in lung tissue. As we have found expression in trachea this apparent discrepancy could be due to a cartilage contamination of the lung tissue from which the libraries were prepared or to that the expression level in lung is too low for a detection with Northern hybridization.

Acknowledgements: This work was supported by Köln Fortune program/Faculty of Medicine, University of Cologne. We thank F. Deák, P. Maurer and D. Newrzella for valuable discussions.

References

[1] J. Engel, V.P. Efimov, and P. Maurer, Development (Camb.) Suppl. (1994) 35–42.
[2] Doolittle, R.F. (1995) Annu. Rev. Biochem. 64, 287–314.
[3] Colombatti, A. and Bonaldo, P. (1991) Blood 77, 2305–2315.
[4] Deák, F., Piecha, D., Bachrati, C., Paulsson, M. and Kiss, I. (1997) J. Biol. Chem. 272, 9268–9274.
[5] Hauser, N. and Paulsson, M. (1994) J. Biol. Chem. 269, 25747–25753.
[6] Paulsson, M. and Heinegård, D. (1979) Biochem. J. 183, 539–545.
[7] Hauser, N., Paulsson, M., Heinegård, D. and Mörgelin, M. (1996) J. Biol. Chem. 271, 32247–32252.
[8] Winterbottom, N., Tondravi, M.M., Harrington, T.L., Klier, F.G., Vertel, B.M. and Goetinck, P.F. (1992) Dev. Dyn. 193, 266–276.
[9] Chen, Q., Johnson, D.M., Haudenschild, D.R., Tondravi, M.M. and Goetinck, P.F. (1995) Mol. Biol. Cell 6, 1743–1753.
[10] Paulsson, M. and Heinegård, D. (1982) Biochem. J. 207, 207–213.
[11] Kiss, I., Deák, F., Holloway, R.G., Delius, H., Mebust, K.A., Frimberger, E., Argraves, W.S., Tsonis, P.A., Winterbottom, N. and Goetinck, P.F. (1989) J. Biol. Chem. 264, 8126–8134.
[12] Jenkins, R.N., Osborne-Lawrence, S.I., Sinclair, A.K., Eddy,

R.L., Byers, M.G., Shows, T.B. and Duby, A.D. (1990) J. Biol. Chem. 265, 19624–19631.
[13] Aszódi, A., Hauser, N., Studer, D., Paulsson, M., Hiripi, L. and Bösze, Z. (1996) Eur. J. Biochem. 236, 970–977.
[14] Lennon, G., Auffray, C., Polymeropoulos, M. and Soares, M.B. (1996) Genomics 33, 151–152.
[15] Genetics Computer Group (1994) Program Manual for the Wisconsin Package, Version 8, Genetics Computer Group, Madison, WI.
[16] Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997) Prot. Eng. 10, 1–6.
[17] Hansen, J.E., Lund, O., Engelbrecht, J., Bohr, H., Nielsen, J.O., Hansen, J.E.S. and Brunak, S. (1995) Biochem. J. 308, 801–813.
[18] Lee, J.-O., Rieu, P., Arnaout, M. and Liddington, R. (1995) Cell 80, 631–638.
[19] Perkins, S.J., Smith, K.F., Williams, S.C., Haris, P.I., Chapman, D. and Sim, R.B. (1994) J. Mol. Biol. 238, 104–119.
[20] Lupas, A., Van Dyke, M. and Stock, J. (1991) Science 252, 1162–1164.
[21] Haudenschild, D.R., Tondravi, M.M., Hofer, U., Chen, Q. and Goetinck, P.F. (1995) J. Biol. Chem. 270, 23150–23154.
[22] Downing, A.K., Knott, V., Werner, J.M., Cardy, C.M., Campbell, I.D. and Handford, P.A. (1996) Cell 85, 597–605.
[23] K. Beck, J.E. Gambee, A. Kamawal, and H.P. Bächinger, EMBO J., 1997 in press.