

# Matrilin-3 from chicken cartilage

Daniele Belluoccio, Beat Trueb\*

MEM Institute, Division of Biology, University of Bern, P.O. Box 30, CH-3010 Bern, Switzerland

Received 19 August 1997

**Abstract** By subtractive cDNA cloning we have identified a novel constituent of chicken cartilage termed matrilin-3. This constituent is encoded by a mRNA of 2.2 kbp whose expression is restricted to cartilaginous tissues. The predicted protein is composed of 452 amino acids with a molecular mass of 49 kDa. It contains a single von Willebrand factor A-like domain, four epidermal growth factor-like repeats and an  $\alpha$ -helical region which may induce the formation of oligomers via a coiled-coil. The primary structure is similar to that of matrilin-1 which is also expressed in a cartilage-specific manner. This similarity suggests that the genes for the two proteins may have evolved from a common ancestor by gene duplication.

© 1997 Federation of European Biochemical Societies.

**Key words:** Cartilage; Epidermal growth factor; Extracellular matrix; Matrilin; Subtractive cDNA cloning; Von Willebrand factor

## 1. Introduction

Alterations in the synthesis and degradation of cartilage components play an important role in the etiology of many degenerative and inherited skeletal disorders, including osteoarthritis, rheumatoid arthritis and several forms of chondrodysplasia. To understand the underlying cause of these alterations, it is inevitable to study the structure and expression of the individual components from normal cartilage. So far only very few cartilage-specific components have been described [1]. This fact might be explained by the tenacious insolubility of most cartilage proteins in physiological buffers. Nevertheless, it has been established that the extracellular matrix of healthy cartilage consists of collagens (collagen II, collagen IX), non-collagenous proteins (matrilin-1 or CMP, cartilage oligomeric matrix protein or COMP) and proteoglycans (aggrecan) [1].

In order to identify additional cartilage-specific components and to overcome the problems caused by their insolubility, we have constructed a subtracted cDNA library with RNA from cartilage and skin. Here we describe one of the clones obtained in this library which encodes a novel cartilage component related to matrilin-1.

## 2. Materials and methods

### 2.1. RNA isolation and Northern blotting

RNA was extracted from various tissues of 16-day-old chicken embryos by the guanidinium-isothiocyanate method [2] utilizing the RNeasy Mini Kit of Qiagen GmbH (Germany). The samples (10  $\mu$ g/

lane) were resolved on 1% agarose gels in the presence of formaldehyde and transferred to nylon membranes by vacuum blotting [3]. The membranes were hybridized overnight under standard conditions (42°C, 50% formamide) with various cDNA probes which had been labeled by the random primed oligolabeling method [4]. The blots were washed at regular stringency and exposed to X-ray film.

### 2.2. Preparation and screening of cDNA libraries

Poly(A)-rich RNA was prepared from total RNA by chromatography on oligo-dT resin using the Oligotex mRNA purification kit (Qiagen). A subtracted cDNA library was generated with poly(A) RNA from chicken sterna and chicken skin by the biotin/streptavidin/phenol method [5,6]. Clone Hb67 derived from this library was used to screen a commercial cDNA library (Clontech, Palo Alto, CA) prepared from 10-day-old chicken embryos by the plaque hybridization technique [7]. Positive clones were picked and subcloned into the plasmids pUC19 and M13mp18 [3].

### 2.3. DNA sequencing

The sequences of the DNA inserts were determined on both strands by the dideoxy chain termination method [8]. The nucleotide and the derived amino acid sequences were compared with all entries of the EMBL and PIR databanks. Alignments were performed with the GCG Computer Program Package (University of Wisconsin, Madison, WI) using default settings.

### 2.4. In situ hybridization

Sterna from 16-day-old chicken embryos were embedded in Tissue-Tek (Sakura Finetek Europe, Zoeterwoude, Netherlands), cut to 10  $\mu$ m sections and placed on microscopy slides. The slides were processed for non-radioactive in situ hybridization following an established protocol [9]. Probes were labeled with digoxigenin-dUTP using the DIG DNA Labeling and Detection Kit (Boehringer Mannheim GmbH, Germany). They included cDNA clones for collagen II (Hb24) and matrilin-3 (Hb67) that had been obtained during this study as well as a cDNA clone for chicken collagen I that had been used previously for in situ hybridization [10]. After mounting, the slides were inspected with an Olympus Vanox-S microscope and photographed.

## 3. Results

### 3.1. A cartilage-specific cDNA library

A subtracted cDNA library was constructed with the mRNA of sterna and skin from chicken embryos. The procedure involved transcription of the two RNA pools into cDNA, hybridization of the two cDNA populations, extraction of hybrids by the streptavidin/phenol method and amplification of the remaining cDNA molecules by PCR. After three rounds of hybridization and subtraction, we obtained more than 100 individual cDNA clones. The inserts of these clones ranged from 200 to 700 bp. When analyzed on Northern blots containing RNA from embryonic sterna and skin, about 50% of the clones showed expression in cartilage but not in skin as expected (Fig. 1, left). Sequencing studies demonstrated that the majority of these clones encoded typical cartilage proteins, including collagen II, collagen IX and aggrecan. In addition, seven cDNA clones were obtained whose sequences were not yet stored in the EMBL databank.

One of these clones (Hb67) was chosen for further inves-

\*Corresponding author. Fax: (41) (31) 632 4999.

**Abbreviations:** CMP, cartilage matrix protein (matrilin-1); EGF domain, epidermal growth factor-like domain; EST, expressed sequence tag; PCR, polymerase chain reaction; vWFA domain, von Willebrand factor A-like domain

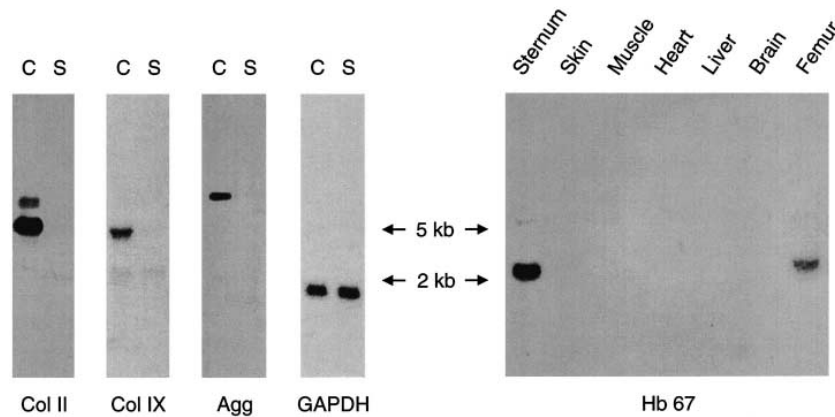


Fig. 1. Northern blot analysis. RNA from eight different tissues as indicated (C, cartilage (sternum); S, skin) was resolved on agarose gels, transferred to nylon membranes and hybridized with radiolabeled probes specific for  $\alpha 1(\text{II})$  collagen (clone Hb24),  $\alpha 1(\text{IX})$  collagen (Hb63), aggrecan (Hb3), matrilin-3 (Hb67) and glyceraldehyde-3-phosphate dehydrogenase [6]. The migration positions of the two ribosomal subunits 18S and 28S are marked.

tigations. When analyzed on a Northern blot with RNA from seven different tissues of chicken embryos, it hybridized to a mRNA of 2.2 kbp that was expressed at relatively high level in the sternum, at moderate level in the femur, but not at all in skin, skeletal muscle, heart, liver and brain (Fig. 1, right). Since femur contains a considerable proportion of cartilage at the developmental stage investigated (E16), we also studied the expression in calvaria, a bone that does not form from cartilage. RNA from calvaria did not hybridize with the Hb67 probe (not shown), suggesting that clone Hb67 showed cartilage-specific expression.

### 3.2. Nucleotide and derived amino acid sequence of clone Hb67

The insert of Hb67 (391 bp) was used as a probe to screen a commercial cDNA library prepared from chicken embryos. Seven overlapping cDNA clones were obtained whose combined nucleotide sequence covered 1836 bp. This sequence started with a 5' untranslated region of 23 nucleotides (Fig. 2). The first ATG codon fulfilled the criteria of a typical translation start site [11]. This codon was followed by an open reading frame of 1356 bp. At the 3' end there was an untranslated region of 457 bp which led into a poly(A) tail. This poly(A) tail was preceded by the polyadenylation signal AA-TAAA at position 1817–1822.

The protein predicted from the open reading frame consisted of 452 amino acids with a molecular weight of 49 622 and an isoelectric point of 5.4 (Fig. 2). It started with a hydrophobic signal sequence which is likely to be cleaved after glycine at position 24 [12]. The mature protein of 428 residues ( $M_r$  47 160) contained one glycosylation signal NGS (position 295–297) which may be utilized for modification by asparagine-linked carbohydrates.

### 3.3. Domain structure and similarity to other proteins

A detailed search of the PIR databank showed that the novel protein was related to two other proteins, matrilin-1 (CMP) and matrilin-2. Compared with chicken matrilin-1 [13], the novel sequence showed 64% identity or 73% similarity if conservative amino acid substitutions were included. An identity of 47% (similarity 57%) was observed with matrilin-2 [14]. The novel protein must therefore belong to the protein family of the matrilins and consequently, it will be termed matrilin-3 in the following.

Matrilin-1 and matrilin-2 are known to have a characteristic multi-domain structure [13–15]. Matrilin-1 consists of two vWFA domains which are connected by an EGF domain. Matrilin-2 contains two vWFA domains that are linked by 10 EGF domains. When the domain structure of the novel protein was analyzed, we found a single vWFA domain at the N-terminus that was followed by four EGF motifs (Fig. 3). The vWFA domain of the novel matrilin showed high similarity to the first vWFA motif (65% identity) and somewhat lower similarity to the second vWFA motif of matrilin-1 (42% identity). The four EGF repeats of matrilin-3 were highly conserved among themselves (68–76% identity) and shared 54–64% identity with the single EGF motif of matrilin-1 (Fig. 3).

Both matrilin-1 and matrilin-2 contain several heptad repeats at their C-terminus forming an  $\alpha$ -helix which enables the polypeptide to assemble into oligomeric structures. By the formation of a coiled-coil, matrilin-1 has been demonstrated to assemble into a trimeric, bouquet-like structure which is stabilized by three interchain disulfide bonds [16,17]. The C-terminal region of matrilin-3 was found to be structurally related to the coiled-coil region of matrilin-1 (26% identity, 46% similarity). All five heptad repeats with their characteristic distribution of hydrophobic and charged residues as well as the positions of the cysteine residues were fully conserved between the two proteins (not shown). It is therefore likely that matrilin-3 will also assemble into an oligomeric structure by the formation of a coiled-coil as depicted in Fig. 4. However, the exact oligomerization state (trimer, tetramer or pentamer) remains to be established.

### 3.4. Human matrilin-3

A search through the databank of expressed human sequence tags yielded one clone (EST T94707) that was highly similar to the chicken sequence. When the full nucleotide sequence of this clone was determined, we realized that it encoded about 60% of the homologous human matrilin-3 sequence (Fig. 2). The derived (partial) amino acid sequence started within the vWFA domain and extended all the way to the C-terminus. At the nucleotide level, the human sequence shared 67% identity with the chicken sequence. At the amino acid level, the identity was 62% (similarity 71%). The modular structures of the two proteins were fully con-

Fig. 2. Complete nucleotide and derived amino acid sequence of chicken matrilin-3 (databank accession number AJ000055). The signal peptidase cleavage site is indicated by an arrow, cysteines are encircled, a potential glycosylation site is boxed. The partial amino acid sequence of human matrilin-3 as derived from the nucleotide sequence of EST clone T94707 (databank accession number AJ001047) is included below the chicken amino acid sequence.

been present during the preparation of the lung-specific cDNA library.

To gain information about the expression of matrilin-3 in cartilage, we analyzed cross-sections of embryonic chicken sterna by *in situ* hybridization (Fig. 5). The mRNA for ma-

M3-vWFA1	45	D	T	A	G	K	N	R	P	L	D	L	V	F	I	D	S	S	R	S	V	R	P	E	E	F	E	K	V	K	I	F	L	77	
M1-vWFA1	30	G	T	L	C	R	T	K	P	T	D	L	V	F	I	D	S	S	R	S	V	R	P	Q	E	F	E	K	V	K	V	F	L	62	
M1-vWFA2	263	C	S	G	G	S	G	S	A	L	D	L	V	F	L	I	D	G	S	K	S	V	R	P	E	N	F	E	L	V	K	K	F	I	295
M3-vWFA1	78	S	K	M	I	D	T	L	D	V	G	E	R	T	T	R	V	A	V	M	N	Y	A	S	T	V	K	V	E	F	P	L	R	T	110
M1-vWFA1	63	S	R	V	I	E	G	L	D	V	G	P	N	S	T	R	V	G	V	I	N	Y	A	S	A	V	K	N	E	F	S	L	K	T	95
M1-vWFA2	296	N	Q	I	V	E	S	L	E	V	S	E	K	Q	A	Q	V	G	L	V	Q	Y	S	S	S	V	R	Q	E	F	P	L	G	Q	328
M3-vWFA1	111	Y	F	D	K	A	S	M	K	E	A	V	S	R	I	Q	P	L	S	A	G	T	M	T	G	L	A	I	Q	A	A	M	D	E	143
M1-vWFA1	96	H	Q	T	K	A	E	L	L	Q	A	V	Q	R	I	E	P	L	S	T	G	T	M	T	G	L	A	I	Q	F	A	I	S	R	128
M1-vWFA2	329	F	K	N	K	K	D	I	K	A	A	V	K	K	M	A	Y	M	E	K	G	T	M	T	G	Q	A	L	K	Y	L	V	D	S	361
M3-vWFA1	144	V	F	T	E	E	M	G	T	R	P	A	N	F	N	I	P	K	V	V	I	I	V	T	D	G	R	P	Q	D	Q	V	E	N	176
M1-vWFA1	129	A	F	S	D	T	E	G	A	R	L	R	S	P	N	I	N	K	V	A	I	V	V	T	D	G	R	P	Q	D	G	V	Q	D	161
M1-vWFA2	362	S	F	S	I	A	N	G	A	R	P	-	-	-	G	V	P	K	V	G	I	V	F	T	D	G	R	S	Q	D	Y	I	T	D	391
M3-vWFA1	177	V	A	A	N	A	R	T	A	G	I	E	I	Y	A	V	G	V	G	R	A	D	M	Q	S	L	R	I	M	A	S	E	P	L	209
M1-vWFA1	162	V	S	A	R	A	R	Q	A	G	I	E	I	F	A	I	G	V	G	R	V	D	M	H	T	L	R	Q	I	A	S	E	P	L	194
M1-vWFA2	392	A	A	K	K	A	K	D	L	G	F	R	M	F	A	V	G	V	G	N	A	V	E	D	E	L	R	E	I	A	S	E	P	V	424
M3-vWFA1	210	D	E	H	V	F	Y	V	E	T	Y	G	V	I	E	K	L	T	S	K	F	R	E	T	F	C								234	
M1-vWFA1	195	D	D	H	V	D	Y	V	E	S	Y	S	V	I	E	K	L	T	H	K	F	Q	E	A	F	C								219	
M1-vWFA2	425	A	E	H	Y	F	Y	T	A	D	F	R	T	I	S	N	I	G	K	K	L	Q	M	K	I	C								449	
M3-EGF2	276	A	V	D	V	C	A	P	G	R	H	E	C	D	Q	I	C	V	S	N	N	G	S	Y	V	C	E	C	F	E	G	Y	T	L	308
M3-EGF3	317	A	M	D	V	C	A	P	G	R	H	D	C	A	Q	V	C	R	R	N	G	S	Y	S	C	D	C	F	E	G	F	T	L	349	
M3-EGF4	358	A	V	D	V	C	A	P	G	R	H	D	C	E	Q	V	C	V	R	D	D	L	F	Y	T	C	D	C	Y	Q	G	Y	V	L	390
M3-EGF1	235	A	A	N	T	C	A	L	G	T	H	D	C	E	Q	V	C	V	S	N	D	G	S	Y	L	C	D	C	Y	E	G	Y	T	L	267
M1-EGF1	221	V	S	D	L	C	A	T	G	D	H	D	C	E	Q	I	C	I	S	T	P	G	S	Y	K	C	A	C	K	E	G	F	T	L	253
M3-EGF2	309	N	P	D	K	K	T	C	S																									318	
M3-EGF3	350	N	P	D	K	K	T	C	S																									357	
M3-EGF4	391	N	P	D	K	K	T	C	S																									398	
M3-EGF1	268	N	P	D	K	R	T	C	S																									275	
M1-EGF1	254	N	N	D	G	K	T	C	S																									282	

Fig. 3. Alignment of the vWFA domains and the EGF repeats from chicken matrilin-1 (M1) and chicken matrilin-3 (M3). Identical residues are boxed.

trilin-3 was detected in chondrocytes throughout the sternal tissue. The signal obtained with our probe was more pronounced at the periphery than in the center of cartilage, suggesting that cells at the surface of the sternum were synthetically more active. The expression pattern was very similar to that obtained with a probe for collagen II. The mRNA for collagen II was also detected throughout the sternum and the signal was also more intense at the periphery of cartilage. A probe for collagen I, which was included as control, revealed a very different distribution. Expression of the mRNA for collagen I was found in the perichondrium and in the connective tissue surrounding the sternum, but not in cartilage.

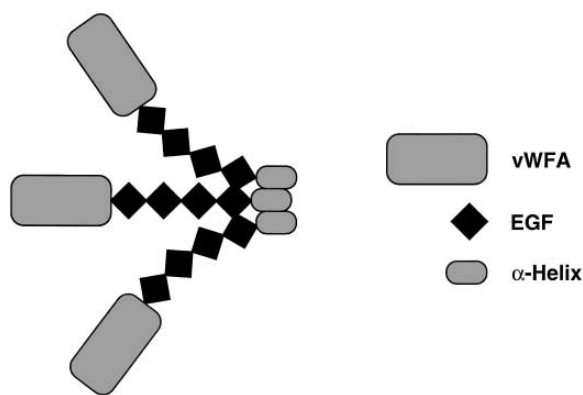


Fig. 4. Domain structure of chicken matrilin-3. A putative assembly of three identical subunits is depicted.

#### 4. Discussion

By subtractive cDNA cloning we have identified a novel component of chicken sternum termed matrilin-3. The expression of the corresponding gene is restricted to cartilage similar to that of collagen II, collagen IX and aggrecan. Matrilin-3 exhibits a modular structure composed of one vWFA domain, four EGF repeats and an  $\alpha$ -helical motif. The overall structure is therefore similar to that of matrilin-1 [13] and matrilin-2 [14], yet the number of domains and their arrangement clearly differ among the three members of this novel protein family. Matrilin-3 was found to be most closely related to matrilin-1, whose expression is also restricted to cartilage. It is therefore possible that the genes for the two proteins evolved from a common ancestor by gene duplication. Compared to matrilin-1, however, the second vWFA domain is lacking in matrilin-3 and the EGF motifs are directly linked to the coiled-coil region.

The function of the matrilins is not yet known in detail. However, the distribution of the proteins, their three-dimensional structure and their interaction with other matrix components may provide valuable clues about possible functions. Our in situ hybridization studies revealed that the mRNA for matrilin-3 co-localizes with that of collagen II, at least in chicken sternum. The sequence similarity with matrilin-1 which is known to form a trimeric, bouquet-like structure [16,17] suggests that matrilin-3 may also form a higher ordered structure. Tentatively, we have drawn such an oligomeric structure in Fig. 4 with three subunits. It must be noted,

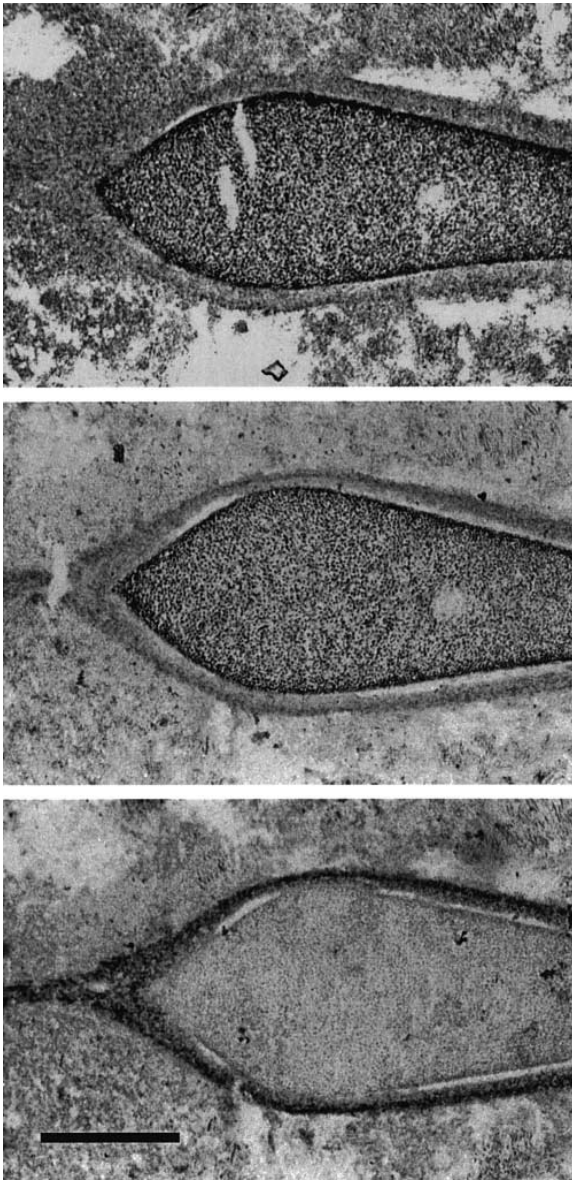


Fig. 5. In situ hybridization studies with chicken sterna. The distribution of the mRNAs for matrilin-3 (top), collagen II (middle) and collagen I (bottom) is presented. Bar, 500  $\mu$ m.

however, that the number of subunits in the final assembly is not yet known and that tetrameric as well as pentameric structures would also be possible [18]. Studies with recombinant proteins are now in progress to solve the oligomerization state of chicken matrilin-3.

Matrilin-1 is known to interact with collagen II [19] and with aggrecan [20]. The interaction with collagen II may be exerted by the vWFA domains since vWFA domains of other proteins, including von Willebrand factor, some integrins and collagen VI, are known to bind to fibrillar collagens [21]. An attractive hypothesis would therefore be that matrilin-3 asso-

ciates with type II collagen fibrils via its vWFA domains. Since an oligomer of matrilin-3 contains several vWFA domains, it could interact at the same time with several fibrils and hold them together with a distance dictated by the number of the EGF repeats. This hypothesis could now be tested in vitro by analyzing the interaction of recombinant matrilin-3 (and of fragments derived therefrom) with collagen II.

Sequence comparisons have also allowed the identification of a cDNA clone for the human homologue of chicken matrilin-3. This clone might be helpful in the localization of the corresponding gene in the human genome. Given the broad distribution of matrilin-3 in all types of cartilage investigated so far, one may expect that mutations in this gene may cause a phenotype resembling that of mutations in other cartilage-specific genes [22]. This possibility could now be tackled by genetic linkage analyses.

## References

- [1] Ratcliffe, A. and Mow, V.C. (1996) in: *Extracellular Matrix* (Comper, W.D., Ed.), Vol. 1, pp. 234–302, Harwood Academic, Amsterdam.
- [2] Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* 162, 156–159.
- [3] Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1987) *Current Protocols in Molecular Biology*, Greene, New York.
- [4] Feinberg, A. and Vogelstein, B. (1983) *Anal. Biochem.* 132, 6–13.
- [5] Duguid, J.R. and Dinauer, M.C. (1990) *Nucleic Acids Res.* 18, 2789–2792.
- [6] Schenker, T., Lach, C., Kessler, B., Calderara, S. and Trueb, B. (1994) *J. Biol. Chem.* 269, 25447–25453.
- [7] Benton, W.D. and Davis, R.W. (1977) *Science* 196, 180–182.
- [8] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- [9] *Nonradioactive In Situ Hybridization, Application Manual*, 2nd edn. (1996) Boehringer Mannheim GmbH, Biochemica, Mannheim.
- [10] Wälchli, C., Koch, M., Chiquet, M., Odermatt, B.F. and Trueb, B. (1994) *J. Cell Sci.* 107, 669–681.
- [11] Kozak, M. (1987) *Nucleic Acids Res.* 15, 8125–8148.
- [12] Von Heijne, G. (1986) *Nucleic Acids Res.* 14, 4683–4690.
- [13] Argraves, W.S., Deak, F., Sparks, K.J., Kiss, I. and Goetinck, P.F. (1987) *Proc. Natl. Acad. Sci. USA* 84, 464–468.
- [14] Deak, F., Piecha, D., Bachrati, C., Paulsson, M. and Kiss, I. (1997) *J. Biol. Chem.* 272, 9268–9274.
- [15] Bork, P. (1992) *Curr. Opin. Struct. Biol.* 2, 413–421.
- [16] Beck, K., Gambee, J.E., Bohan, C.A. and Bächinger, H.P. (1996) *J. Mol. Biol.* 256, 909–923.
- [17] Haudenschild, D.R., Tondravi, M.M., Hofer, U., Chen, Q. and Goetinck, P.F. (1995) *J. Biol. Chem.* 270, 23150–23154.
- [18] Beck, K., Gambee, J.E., Kamawal, A. and Bächinger, H.P. (1997) *EMBO J.* 16, 3767–3777.
- [19] Winterbottom, N., Tondravi, M.M., Harrington, T.L., Klier, F.G., Vertel, B.M. and Goetinck, P.F. (1992) *Dev. Dyn.* 193, 266–276.
- [20] Hauser, N., Paulsson, M., Heinegard, D. and Mörgelin, M. (1996) *J. Biol. Chem.* 271, 32247–32252.
- [21] Colombatti, A., Bonaldi, P. and Doliana, R. (1993) *Matrix* 13, 297–306.
- [22] Williams, C.J. and Jimenez, S.A. (1995) *J. Rheumatol. (Suppl. 43)* 22, 28–33.