Identification of a Matrix-binding Domain in MAGP1 and MAGP2 and Intracellular Localization of Alternative Splice Forms*

Received for publication, October 28, 2001, and in revised form, January 14, 2002 Published, JBC Papers in Press, January 16, 2002, DOI 10.1074/jbc.M110347200

Fernando Segade‡, Barbara Crippes Trask‡, Thomas J. Broekelmann‡, Richard A. Pierce¶, and Robert P. Mecham‡¶

From the ‡Department of Cell Biology and Physiology and the ¶Division of Pulmonary and Critical Care Medicine, Department of Medicine, Barnes-Jewish Hospital and Washington University School of Medicine, St. Louis, Missouri 63110

MAGP1 is a small molecular mass protein associated with microfibrils in the extracellular matrix (ECM). To identify the molecular basis of its interaction with other microfibrillar proteins, deletion constructs of MAGP1 were expressed in a mammalian cell system that served as a model for microfibril assembly. This study identified a 54-amino acid sequence in the carboxyl-terminal region of the protein that defines a matrix-binding domain that is sufficient to target MAGP1 to the ECM. Site-directed mutagenesis demonstrated that binding activity is dependent on the presence of 7 cysteine residues in this sequence. MAGP2 contains a sequence similar to the matrix-binding domain of MAGP1, but could not associate with the ECM because of a single amino acid change. Two naturally occurring MAGP1 splice variants, MAGP1B (human-specific) and MAGP1D (found in mice), localized intracellularly when expressed as chimeric proteins with green fluorescent protein in rat lung fibroblasts. This suggests a second action site for MAGP1.

Microfibrils are 10–12-nm beaded filaments found in most tissues, particularly elastic tissues, bone, and cartilage (1). Their best known function is as a scaffold for tropoelastin deposition in the assembly of elastic fibers, although there is increasing evidence that they contribute to the mechanical properties of elastic tissues, provide structural integrity to non-elastic tissues, and anchor cells to elastic fibers and other matrix components through cell adhesion motifs (2). The major structural proteins of the microfibril are the fibrillins (3, 4), large modular glycoproteins consisting of repeating calcium-binding epidermal growth factor-like domains interspersed between unique 8-cysteine domains. The fibrillins create the core structure of microfibrils through a head-to-tail alignment of monomer bundles (5). Ultrastructurally, microfibrils are characterized by a “beads on a string” appearance, with fibrillin molecules contributing to both the bead and interbead regions. Other proteins have been localized to microfibrils in many tissues, but their exact function is unknown.

MAGP1 (6) and its structural relative MAGP2 (7) are two of the proteins that associate covalently with fibrillin-containing microfibrils (8). These proteins have no significant similarity to any other proteins in the data bases and therefore form a small unique family of microfibrillar proteins. MAGP1 contains 183 amino acids that form two distinctive domains: the amino-terminal half of the molecule is highly acidic and enriched in proline and contains a clustering of glutamine residues, whereas the carboxyl-terminal portion contains all 13 cysteine residues and has an overall net positive charge. Disulfide bonding between cysteines apparently mediates the binding of MAGP1 to the microfibril because the presence of reducing agents is a requirement for the extraction of MAGP1 from tissues (9). The protein undergoes post-translational transglutamination, sulfation of specific tyrosine residues, and O-linked glycosylation (10). In mouse development, Magp1 mRNA is widely expressed, mainly in mesenchymal and connective tissue cells, where it is easily detected as early as day 8.5 of development (11). MAGP1 is associated with all microfibrils with only two known exceptions: those directly adjacent to the plasma membrane of aortic endothelial cells and those at the junction of the zonule and the lens capsule of the eye (8, 12, 13).

The mammalian MAGP1 gene has been characterized in human (also known as MFAP2) (14), mouse (Magp1) (11), and bovine (15). A comparative study between human and mouse genes and analysis of the public data bases led us to identify novel forms of MAGP1 that arise from the alternative splicing of a unique transcript (16). We identified a total of five splice variants in addition to the canonical full-length MAGP1A form. These transcripts are species-specific and are generated by different processing mechanisms. The alternative splice forms MAGP1A1, MAGP1B, and MAGP1C are expressed in human tissues, while the two variants Magp1A and Magp1D were found only in mouse. The alternatively spliced forms show restricted patterns of expression relative to the canonical isoform MAGP1A. Based on their features, we predicted that human MAGP1B and mouse Magp1D are not extracellular proteins, but represent possible intracellular forms of MAGP1 (16).

It is not yet known what role MAGP1 plays in microfibrillar structure or function. Previous studies have demonstrated that it is capable of binding tropoelastin, an interaction that may influence the linking of elastin monomers to microfibrils for their proper alignment. The acidic amino-terminal half of MAGP1 has been suggested to interact with a positively charged pocket near the C terminus of the tropoelastin mono-

* This work was supported by Grants HL53325, HL62995, and HL61006 from the National Institutes of Health and by a grant from the National Marfan Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡To whom correspondence should be addressed: Dept. of Cell Biology and Physiology, Washington University School of Medicine, Campus Box 8228, 660 S. Euclid Ave., St. Louis, MO 63110. Tel.: 314-362-2254; Fax: 314-362-2252; E-mail: bmecham@cellbiology.wustl.edu.

1 The abbreviations used are: MAGP, microfibril-associated glycoprotein; ECM, extracellular matrix; GFP, green fluorescent protein; EGFP, enhanced green fluorescent protein; MBD, matrix-binding domain.
mer. Anti-peptide antibodies directed against a domain near the N terminus of MAGP1 are able to block the incorporation of tropoelastin into fibers by cultured chondrocytes (17). Finnis and Gibson (18) localized tropoelastin binding, as well as binding to type VI collagen, to a tyrosine-rich sequence near the amino terminus of MAGP1. This sequence overlaps with the region suggested by the antibody studies of Brown-Augsburger et al. (17).

MAGP2 is a 170–173-amino acid protein with similarity to MAGP1 in a central region of 60 amino acids where there is precise alignment of 7 cysteine residues (7). Outside of this region, the MAGP2 molecule is rich in serine and threonine residues and contains an RGD motif through which it binds \( \alpha_5\beta_1 \) integrin (19). The molecule lacks the proline-, glutamine-, and tyrosine-rich sequences characteristic of MAGP1. Although MAGP2 has been localized to both elastin-associated and elastin-free microfibrils in a number of tissues, the protein exhibits patterns of tissue localization and developmental expression that are more restricted than those of MAGP1 (20, 21). These structural and functional differences suggest that MAGP2 has a function related to cell signaling during microfibril assembly and elastogenesis, whereas MAGP1 may be more important for the structural integrity of the microfibril scaffold (20).

In this report, we describe the identification of domains in MAGP1A that target the protein to the ECM in living cells. To this end, we constructed a number of mutants derived from human MAGP1A that were expressed as fusion proteins with GFP as an in vivo reporter for protein localization. The distribution patterns obtained when these fusion proteins were expressed in rat lung cells identified a 54-amino acid sequence that encodes a minimal ECM-binding domain (MBD) located in the C-terminal region of MAGP1A. Through comparative analysis, a similar domain in MAGP2 was identified, but was found not to associate with the ECM of RFL-6 cells. Interestingly, matrix-binding ability could be generated in the MAGP2 sequence through a single amino acid change based on the matrix-binding ability could be generated in the MAGP2 sequence. This end, we constructed a number of mutants derived from human MAGP1A-GFP Is Targeted to the ECM of Lung Fibroblasts—

EXPERIMENTAL PROCEDURES

Cells, Media, and cDNAs—Rat RFL-6 cells (ATCC CCL-192) were purchased from the American Type Culture Collection (Manassas, VA) and were maintained in Kight’s modified Ham’s nutrient mixture F-12 (Ham’s F-12K, Sigma) supplemented with 2.5 g/liter sodium bicarbonate, 20% fetal bovine serum (Invitrogen), 2 mM glutamine (Cell-BankTM/EBI Data Bank and subsequent conceptual translation. Eval-

grams (DNASTAR, Inc., Madison, WI). Coding sequences of vertebrate translation were performed with the LASERGENE package of pro-

genome systems Inc. (St. Louis, MO).

Bioinformatics—Primary sequence input, alignment, and conceptual translation were performed with the LASERGENE package of pro-

grams (DNASTAR, Inc., Madison, WI). Coding sequences of vertebrate MAGP1 and MAGP2 were obtained by BLAST searches of the Gen-

BankTM/EBI Data Bank and subsequent conceptual translation. Eval-

uation of the evolutionary distances between the coding sequences of the human and mouse genes was carried out with SNAP software (Synonymous/Non-synonymous Analysis Program).2

Antibodies—Polyclonal antibodies against bovine recombinant MAGP1 were as described (10). Polyclonal antibodies against mouse recombinant tropoelastin were generated from a His-Tagged fragment of mouse tropoelastin cDNA (exons 6–17) cloned into the pQE-31 vector (QIAGEN, Inc., Valencia, CA) and expressed in Escherichia coli.7 Monoclonal anti-V5 antibodies were purchased from Invitrogen.

Phenylalanine/Valine Constructs and Mutagenesis—Construction of the chimeric GFP fusion proteins was carried out within the pEGFP-N1 vector (CLON-

TECH, Palo Alto, CA).8 Schematics of the deletion mutants are shown in Fig. 3G. For the generation of deletion mutants, the appropriate oligonucleotides were used in high-fidelity PCR reactions that were performed according to the manufac-

turer’s instructions. A summary of the single residue mutants shown is in Table I. The integrity of the constructs was verified by sequencing with the EGFP-N oligonucleotide (CLONTECH) or the appropriate PCR primers. Samples were sequenced at the Washington University Protein and Nucleic Acid Chemistry Laboratory DNA Sequencing Facility. Sequences were extracted and analyzed with ABI Prism software (Applied Biosystems, Foster City, CA). Plasmids for transfections were purified from bacterial cultures with the QIAGEN EasiPure DNA purification kit.

DNA Transfections—RFL-6 cells were seeded at ~60% confluency onto six-well plates containing sterile coverslips. After 24 h, when the cells reached >90% confluency, the medium was replaced with fresh growth medium immediately before transfection, and the cells were transfected with 5 \( \mu \)g of the appropriate plasmid using 5 \( \mu \)l of the LipofectAMINE 2000 transfection reagent (Invitrogen) according to the manufacturer’s instructions. Cells were then incubated for 48 h before visualization of GFP expression.

RESULTS

MAGP1A-GFP Is Targeted to the ECM of Lung Fibroblasts—

The localization of MAGP1A in living cells was studied by transient transfection of rat RFL-6 cells with a chimeric plasmid coding for full-length human MAGP1A or MAGP1A fused to GFP. We found that RFL-6 cells do not express MAGP1 endogenously, as we were unable to detect the expression of the protein with antibodies against MAGP1 (22) (Fig. 1A) or its mRNA by Northern blot analysis or reverse transcription-coupled PCR (data not shown). Distribution patterns of MAGP1A-GFP were visualized by fluorescence microscopy at 48 and 72 h after transfection. These images show that MAGP1A-GFP was present in the ECM at both time points (Fig. 1B). The pattern of MAGP1A-GFP fusion protein expression was similar to that of unmodified MAGP1 (Fig. 1C) or MAGP1 containing a V5 epitope tag (Fig. 1D), indicating that the presence of the GFP molecule had little effect on the ability of MAGP1 to associate

2 Available at hiv-web.lanl.gov.

3 F. Segade, unpublished data.

4 Sequences of the oligonucleotide primers used for the generation of the EGFP fusion proteins, deletions, and point mutations and detailed cloning procedures are available from the authors upon request.
Matrix-binding Sequences in the MAGPs

![Image](https://via.placeholder.com/150)

**Fig. 1. Expression of MAGP1A constructs in RFL-6 cells.** RFL-6 cells were transfected with tagged and untagged versions of MAGP1A, fixed, and subjected to indirect immunofluorescence staining or direct fluorescence visualization. **A,** untransfected RFL-6 cells stained with polyclonal anti-MAGP1 antibodies; **B,** direct fluorescence detection of human MAGP1A-GFP; **C,** immunofluorescence detection of transfected bovine MAGP1A stained with anti-MAGP1 antibodies; **D,** immunofluorescence detection of human MAGP1A-V5 stained with monoclonal anti-V5 antibodies; **E,** endogenous rat tropoelastin in untransfected RFL-6 cells stained with polyclonal anti-mouse antibodies; **F,** rat tropoelastin in RFL-6 cells transfected with human MAGP1A-GFP. Secondary antibodies were conjugated with Alexa 594 or Alexa 488. Images were acquired with a ×40 objective at a final optical magnification of ×400.

Intracellular Localization of MAGP1 Splice Forms—We previously reported the identification and characterization of two alternative transcripts coding for shortened versions of MAGP1 in human and mouse (16). The characteristics of the variants suggested that they might remain as intracellular proteins due to the internal deletion of significant structural features, such as a fragment of the signal peptide that includes the cleavage site. We transiently transfected RFL-6 cells with plasmids encoding GFP-tagged human MAGP1B and mouse MAGP1D to assess whether the proteins are secreted and associate with the matrix. At 48 h after transfection, there was no visible association of the GFP fusion proteins with the ECM (Fig. 2, A and B), although fluorescence indicative of cell transfection and GFP synthesis was clearly observed. In both cases, the chimeric proteins accumulated throughout the cytoplasm and did not localize with the endoplasmic reticulum or Golgi, suggesting that the variants were not secreted. This was confirmed by the absence of soluble human MAGP1B or mouse MAGP1D in the conditioned media from transfected cells as determined by Western blotting (data not shown). These results indicate that the absence of fluorescence associated with the ECM is due to the non-secretion of the alternative splice forms and not to an alteration in their structure that prevents them from binding matrix proteins.

Although the transfection efficiencies of the human MAGP1B-GFP and mouse MAGP1D-GFP plasmids were similar to those obtained with MAGP1A-GFP and pEGFP-N1 (∼25–30% of the cells), we observed that cells expressing the alternative splice forms underwent morphological changes 48 h after transfection. Compared with cells expressing MAGP1A-GFP (Fig. 2A) or EGFP (Fig. 2D), human MAGP1B-GFP and mouse MAGP1D-GFP transfectants showed shrinkage of their cell membranes with retraction of the cytoplasm and a gradual decrease in size. Many cells contained vacuolar inclusions; and in those that were most probably dead and about to detach from the substrate, nuclear localization of the fusion protein was observed. At 4 days after transfection, most of the expressing cells had died and detached from the coverslips. We never observed a significant mortality in cells transfected with MAGP1A-GFP (Fig. 2A) or EGFP-N1 (Fig. 2D) or with any of the other mutants we generated (see below). Therefore, a general effect on cell health of the overexpression of exogenous proteins can be safely ruled out. We speculate that cell mortality may be related to the rate that the alternative forms play in vivo.

Mapping of the ECM-binding Domain of MAGP1A—Although a number of proteins have been shown to interact with MAGP1A (18, 22, 23), the critical domain responsible for the association of MAGP1 with the ECM has not been functionally defined. During our analysis of the coding sequence of mammalian MAGP1, we classified its exons into two classes according to their rate of non-synonymous (non-silent) substitution of codons (16): exons 3, 4, 7, and 8 are evolving more slowly than the average rate for the full sequence, whereas the remaining exons undergo sequence substitution at a much higher rate. Because a slow substitution rate of codons is usually related to the presence of selective constraints acting on the protein sequence, we focused our efforts on the slowly evolving exons in MAGP1 to specifically define their role in ECM binding. To test our hypothesis that two functional units might be present in exons 3 and 4 and exons 7 and 8, a series of deletion mutants derived from human MAGP1A fused to GFP were transfected into RFL-6 cells. The first deletion mutant, MAGP1ΔEx9-GFP (Fig. 3B), a C-terminal truncation of residues 151–183 encoded by exon 9, associated with the ECM, similar to the full-length protein (Fig. 3A). However, mutant MAGP1ΔEx7–9-GFP, in which the entire cysteine-rich domain of MAGP1A (amino acids...
96–183) was deleted, was not visibly associated with the ECM (Fig. 3C), even though its secretion was not impaired by the deletion (data not shown). These results indicate that exon 9, coding for 7 of the 13 cysteine residues in MAGP1A, is not required for MAGP1 deposition onto the ECM. Similarly, the conserved region encoded by exons 3 and 4 that is reported to mediate the conversion of MAGP1A monomers into higher molecular mass complexes (10) is not by itself capable of mediating the initial interaction of MAGP1 with the ECM.

To test whether exons 7 and 8 are critical for the matrix deposition of the fusion protein, we constructed the MAGP1Ex(7–9)-GFP plasmid, containing the signal sequence fused to residues 96–150 of MAGP1A and GFP. Transfection of RFL-6 cells showed that MAGP1Ex(7+8)-GFP was deposited onto the matrix (Fig. 3D) and that the fluorescence levels associated with the matrix were consistently higher with MAGP1Ex(7+8)-GFP than those obtained with full-length MAGP1A-GFP. Moreover, the sequence encoded in exons 7 and 8 forms a functional unit because separate expression of exon 7 or 8 as a fusion protein with GFP (MAGP1Ex7-GFP or MAGP1Ex8-GFP, respectively) completely abolished the ECM-binding ability showed by MAGP1Ex(7+8)-GFP (Fig. 3, E and F). Taken together, our results identify a region (which we named the matrix-binding domain) in the C-terminal region of
MAGP1A, translated from exons 7 and 8, that is critical and sufficient to direct the deposition of GFP onto the matrix (Fig. 3G). We believe that the MBD may represent the domain responsible for the initial interaction of MAGP1A with the ECM.

Cysteine Residues in the MAGP1 MBD Are Critical for ECM Deposition—The primary structure of the MAGP1 MBD contains 7 cysteine and 2 histidine residues (Fig. 4a), reminiscent of the zinc-binding cysteine-rich motifs involved in protein-protein interaction (24). To identify the critical amino acids in the MBD that are responsible for ECM binding, particularly the cysteine and histidine residues, we constructed a battery of single residue mutants in which the targeted amino acids were substituted with alanine residues in the context of construct MAGP1Ex(7/8)-GFP. The results obtained with this series of mutants are summarized in Fig. 5. The mutagenesis of any of the 6 cysteines in the MBD entirely abrogated the deposition of the fusion protein onto the ECM. Similar results were obtained when 2 nearest-neighbor cysteines were simultaneously mutated (data not shown). We would expect that the substitution of 1 of the histidine amino acids would also cancel the association of the MBD with the ECM if those residues were involved in the coordination of a metal atom critical for function. We found, however, that the H111A and H144A mutants were equally efficient as the wild-type MBD construct in associating with the matrix (Fig. 5). Our data suggest that the functionality of the domain is dependent on the presence of the cysteines, possibly through the formation of inter- and intramolecular disulfide bridges that ensure the proper spatial conformation.

A Single Amino Acid Variation Explains the Non-functionality of the MAGP2 Paralogous MBD—MAGP2 is the second member of the MAGP family and contains a sequence that is similar to the MAGP1 MBD (25). The region from exons 8 and...
Matrix-binding Sequences in the MAGPs

9 in MAGP2 (paralogous to exons 7 and 8 in MAGP1) contains 6 cysteines. At a position equivalent to cysteine 114 of human MAGP2 (valine 101) (Fig. 4A). From our mutagenesis studies, we predicted that MAGP2 should be inactive in an ECM association assay in RFL-6 cells because a valine substituting for cysteine at that position is equivalent to our engineered inactive C114A mutant (Fig. 5). In transfected RFL-6 cells, the GFP fusion proteins containing full-length MAGP2 (MAGP2-GFP) or the putative MAGP2 MBD (MAGP2Ex(8+9)-GFP) were not deposited onto the ECM (Fig. 6, A and B), even at 72 h after transfection (data not shown). If valine 101 were responsible for the loss of ECM-binding activity of the domain, we speculated that changing Val101 to Cys should be able to restore its function. When the mutant protein MAGP2Ex(8+9)V101C-GFP was expressed in RFL-6 cells, it regained the ability to associate with the ECM (Fig. 6C), with a pattern and efficiency that were comparable to those obtained with the MAGP1 MBD construct MAGP1Ex(7+8)-GFP.

**DISCUSSION**

We have previously reported that mammals possess a total of six variants of MAGP1 transcribed from the single copy MAGP1 gene (16). To study the fate of MAGP1 and its alternatively spliced forms, we generated chimeric proteins using the coding sequence of the appropriate variant fused to the coding sequence for GFP. As a suitable in vitro system to assay the subcellular localization, we chose the rat lung fibroblast cell line RFL-6, which we had characterized for the expression of microfibrillar components and their assembly into bona fide microfibrils (26). Because RFL-6 cells do not make endogenous MAGP1, they are optimal for the study of matrix association of heterologous MAGP1 and its derivatives. As we expected, the expression of MAGP1A in any of its tagged and untagged versions did not negatively affect the deposition of other microfibrillar components, such as tropoelastin, with which MAGP1A co-localizes in the ECM. In fact, the fibrillar pattern of MAGP1A-GFP in the ECM of RFL-6 cells is identical to the morphology of endogenous MAGP1A in pigmented epithelium cells or fetal bovine chondrocytes (27). It is important to note that although the efficiency of transfection of RFL-6 cells with our chimeric constructs was estimated at 25–30%, we did not observe the formation of clumps of fluorescent fibrils corresponding to ECM-associated GFP chimeras. Instead, the decoration of fibrils was widespread throughout the cultures, as if transfected cells secrete the protein into the growth medium, where it is later deposited onto the matrix without the intervention of the living cell. A similar phenomenon has been observed for the heterologous expression of tropoelastin in pigmented epithelium cells.5

The expression of human MAGP1B-GFP and mouse MAGP1D-GFP fusion proteins in RFL-6 cells allowed us to visualize the subcellular compartment to which they localize. As we predicted, both proteins were retained inside the cell in a diffuse cytoplasmic pattern (Fig. 2), and they were not deposited onto the ECM. There were no soluble forms of the fusion proteins indicative of secretion, thereby confirming that the alternatively spliced forms are intracellular versions of MAGP1A. In RFL-6 cells and other cell lines transfected with the splice variants, we observed that the expression of human MAGP1B or mouse MAGP1D is a transient event because transfected cells experience alterations in their cell membranes, stop dividing, and die after ~4 days in culture; therefore, no stable transfectants could be obtained. It is possible that the toxicity is induced simply by the overexpression of protein that is normally present in the cell (and only in certain tissues) in a low number of copies, as estimated from transcript abundance (16). Another possibility is that the MAGP1 variants arise from transcripts that are aberrantly spliced, perhaps due to the actions of the repetitive elements inserted around exon 3 in the MAGP1 gene (16, 28). If this is the case, the

---

5 B. Kozel, personal communication.
variant transcripts are not to be translated, but rather degraded to prevent the synthesis of potentially harmful proteins (29). According to this interpretation, the fact that different species have distinct alternative mRNAs that are not phylogenetically conserved might be an indication of their aberrant character. If human MAGP1B and mouse MAGP1D represent functional intracellular forms of MAGP1, what would their role be in an intracellular environment? A semiquantitative study of the expression of the alternative transcripts in tissues showed a more restricted pattern of expression at very low ratios relative to the canonical MAGP1A mRNA (16), suggestive of a tight regulation of the generation of alternative isoforms. Interestingly, MAGP1 contains a cysteine-rich domain with a high similarity to motifs involved in protein-protein interaction present usually in regulatory modulators (see below). In this sense, the observed toxic effects in transfected cells could reflect the endogenous role of intracellular MAGP1, perhaps in tissues undergoing remodeling in which programmed cell death is extensive. In this regard, MAGp1D is expressed in the mouse mammary gland only during involution of the gland (16), a well-characterized process that involves apoptosis (30).

MAGP expression has some similarities to the matricellular protein SPARC. SPARC is a multifunctional glycoprotein that modulates cellular interaction with the ECM, contributing to a counter-adhesive effect on cells (31). Most important, SPARC is expressed at significant levels in remodeling tissues. Challenging the concept that SPARC is exclusively secreted, Gooden et al. (32) have shown the presence of SPARC in the nucleus of dividing cells. Immunoreactive SPARC was found in the nuclei of day 2 chick embryonic cells and adult bovine aortic endothelial cells. The function of SPARC in the nucleus is currently unknown, but the evidence thus far indicates that it might be involved in the regulation of mitosis in pre-differentiated cells (31). Similarly, it is conceivable that MAGP1 is also a bifunctional protein, with a structural microfibril-dependent extracellular role and a regulatory intracellular function.

The MBD of MAGP1 is 54 amino acids in size and characterized by the presence of 7 cysteine residues and an almost equal number of acidic and basic amino acids (9 and 7, respectively). The cysteines in the C terminus of MAGP1 and MAGP2 suggest the formation of a highly structured domain. A number of microfibrillar proteins are also rich in cysteines, including the fibrillins and latent TGF-β-binding proteins (2), which are composed of a large number of repeats of cysteine-stabilized motifs that give the molecules their rigidity and fibrillar conformation. However, the MAGP1 MBD does not conform to the consensus for any known cysteine-rich repeat, and it may well represent a unique invention of the MAGPs (Fig. 4A). Our transfection studies indicated that the MBD is strictly required for the deposition of MAGP1A onto the ECM and that it is the minimal functional unit for binding (Fig. 3). The latter point is especially important in the elucidation of the evolutionary origin of the MAGPs (see below). Single and double residue mutants of the MBD identified every cysteine as critical for maintaining functionality, whereas the substitution of other residues, such as the 2 histidines, had no effect on the affinity of the fusion proteins for the ECM.

In the RFL-6 cell system, the MAGP2-GFP fusion protein was expressed and secreted into the conditioned medium, but was unable to associate with the ECM (Fig. 6). Although this result was somewhat paradoxical given the tight covalent association of both MAGP1A and MAGP2 with the microfibrils (6, 7, 9), we decided to use the paralogous MBD in MAGP2 as an inactive but naturally occurring mutant of the MAGP1A MBD. We reasoned that because cysteines are critical for binding, the substitution of 1 cysteine for a valine residue in the MAGP2 sequence was responsible for the loss of binding shown by MAGP2. The fact that we were able to revert the loss of function by mutating valine 101 to cysteine in the MAGP2Ex(8+9)V101C-GFP chimera clearly confirmed our hypothesis and suggested that the multiple substitutions of amino acids besides Val101 between the MAGP1 and MAGP2 MBDs are essentially neutral in their action. The presence of an odd number of cysteines in the MAGP1 MBD suggests that cysteine 114 might be involved in the formation of a disulfide bond between MAGP1 and its ligand in the ECM. The presence of valine 101 in the MAGP2 MBD would explain the inability of MAGP2 to deposit onto the RFL-6 ECM.

In the previous discussion, we assumed that binding to the ECM is the original function of the MAGP1 MBD, which was subsequently lost during the evolution of MAGP2. Therefore, MAGP1 would be closer to the ancestral gene function than MAGP2. There are several findings that support this view. First, as shown in Fig. 4A, an extensive search of the public domain data bases found MAGP1 sequences in every vertebrate group from teleost fishes through amphibians to mammals, although no sequences corresponding to MAGP1-related genes were found in the nematode and insect data bases. MAGP2 seems to have arisen recently in evolutionary terms because during careful searches of the data bases, only mammalian sequences were identified. Second, MAGP1 is broadly expressed in tissues at all stages of development (11, 16), whereas MAGP2 gene expression is restricted in its temporal and spatial patterns (20, 21). Third, the calculated rate of base substitutions that change the amino acid sequence of the encoded protein (non-synonymous substitution) and the rate of nucleotide substitutions that are silent at the amino acid level (synonymous substitution) for the mammalian MAGP1 and MAGP2 genes showed that the synonymous rate, a measure of the basal mutation rate (33), is essentially similar for both genes (Fig. 4B). However, the MAGP1 genes show a 50% lower level of non-synonymous substitution compared with the MAGP2 genes. In fact, some domains in MAGP2, such as the N-terminal region, are apparently undergoing positive selection for sequence change (non-synonymous rate > synonymous rate). In MAGP1, the sequences from exons 3, 4, 7, and 8 correspond to the horizontal stretches in the non-synonymous rate graph, indicative of absolute conservation of sequence (Fig. 4B). Interestingly, in MAGP2, the lowest non-synonymous rate is shown by exons 8 and 9, paralogous to MAGP1 exons 7 and 8, which we identified as the MAGP1 MBD. In an analysis of vertebrate gene families, Hastings (34) found that the most conserved member of a family is the most broadly expressed one and is the most similar in structure and function to the ancestral gene. In the MAGP family, MAGP1 is the broadly expressed and conserved gene. Because the majority of the protein components of the ECM are modular and conceivably evolved by repeated rounds of exon shuffling (35), we speculate that the ancestral MAGP1 protein contained a functional MBD capable of association with the ECM. Through exon shuffling, the exons coding for the MBD were transposed to a different genomic region, from which the MAGP2 gene evolved. It is important to note that most of the modules involved in exon shuffling are flanked by splice sites of class 1-1 (splitting a codon between the first and second nucleotides) (35). Remarkably, the MBD in MAGP1 and MAGP2 is coded by two exons with splice junctions of classes 1-2 and 2-1, respectively, but because both exons work as a contiguous functional unit, they result in a domain flanked by class 1-1 sites. Once the MBD was duplicated, the ancestral MAGP2 gene diverged enough to lose the ancestral binding activity (and to gain a new specificity), whereas stabilizing selection kept the sequence of MAGP1
from diverging to preserve the essential function performed in a variety of tissues.

Our results demonstrate that the inability of MAGP2 to associate with the ECM of RFL-6 cells is due to the divergent sequence in the paralogous MBD and that the loss of function can be restored by a single residue mutation. This finding has important implications not only for understanding the evolutionary origin of the MAGPs, but also for elucidating the functional specificity of the closely related MBDs in MAGP1 and MAGP2 for their interaction with the microfibrillar complex. Interactions of MAGP1 with tropoelastin (23), fibrillin-1 (22), and fibrillin-2 (26) have been described. Although deposited onto the RFL-6 ECM, tropoelastin is an unlikely candidate for the MAGP1 binding partner because the interaction between these two molecules has been mapped to an N-terminal acidic region of MAGP1 (17) that does not correspond to the MBD identified here. However, the fibrillins represent a more likely target for the MAGPs. In fact, in ligand blot assays, full-length bovine recombinant MAGP1 associates with an N-terminal fragment of fibrillin-2 (26). However, preliminary studies on the interaction between fibrillins and MAGP1 using the yeast two-hybrid assay system seem to indicate that although both MAGP1 and MAGP2 MBDs are capable of association with fibrillin-2, they do not show the specificity observed in the association with the RFL-6 ECM. Experiments to identify the binding partner of the MAGP1 MBD are currently under way in our laboratory.

Acknowledgments—We thank Clarina Tisdale and Sarah Broekelmann for excellent technical support.

REFERENCES
1. Mecham, R. P., and Davis, E. C. (1994) in Extracellular Matrix Assembly and Structure (Vurchenko, P. D., and Mecham, R. P., eds) pp. 281–314, Academic Press, Inc., San Diego, CA
2. Robinson, P. N., and Godfrey, M. (2000) J. Med. Genet. 37, 9–25
3. Sakai, L. Y., Keene, D. R., and Engvall, E. (1986) J. Cell Biol. 103, 2499–2509
4. Zhang, H., Apfelroth, S. D., Hu, W., Davis, E. C., Sanguineti, C., Bonadio, J., Mecham, R. P., and Ramirez, F. (1994) J. Cell Biol. 124, 855–863
5. Reinhardt, D. P., Keene, D. R., Corson, G. M., Poschl, E., Bachinger, H. P., Gambee, J. E., and Sakai, L. Y. (1996) J. Mol. Biol. 256, 104–116
6. F. Segade and C. Werneck, unpublished data.

from diverging to preserve the essential function performed in a variety of tissues.

Our results demonstrate that the inability of MAGP2 to associate with the ECM of RFL-6 cells is due to the divergent sequence in the paralogous MBD and that the loss of function can be restored by a single residue mutation. This finding has important implications not only for understanding the evolutionary origin of the MAGPs, but also for elucidating the functional specificity of the closely related MBDs in MAGP1 and MAGP2 for their interaction with the microfibrillar complex. Interactions of MAGP1 with tropoelastin (23), fibrillin-1 (22), and fibrillin-2 (26) have been described. Although deposited onto the RFL-6 ECM, tropoelastin is an unlikely candidate for the MAGP1 binding partner because the interaction between these two molecules has been mapped to an N-terminal acidic region of MAGP1 (17) that does not correspond to the MBD identified here. However, the fibrillins represent a more likely target for the MAGPs. In fact, in ligand blot assays, full-length bovine recombinant MAGP1 associates with an N-terminal fragment of fibrillin-2 (26). However, preliminary studies on the interaction between fibrillins and MAGP1 using the yeast two-hybrid assay system seem to indicate that although both MAGP1 and MAGP2 MBDs are capable of association with fibrillin-2, they do not show the specificity observed in the association with the RFL-6 ECM. Experiments to identify the binding partner of the MAGP1 MBD are currently under way in our laboratory.

Acknowledgments—We thank Clarina Tisdale and Sarah Broekelmann for excellent technical support.

REFERENCES
1. Mecham, R. P., and Davis, E. C. (1994) in Extracellular Matrix Assembly and Structure (Vurchenko, P. D., and Mecham, R. P., eds) pp. 281–314, Academic Press, Inc., San Diego, CA
2. Robinson, P. N., and Godfrey, M. (2000) J. Med. Genet. 37, 9–25
3. Sakai, L. Y., Keene, D. R., and Engvall, E. (1986) J. Cell Biol. 103, 2499–2509
4. Zhang, H., Apfelroth, S. D., Hu, W., Davis, E. C., Sanguineti, C., Bonadio, J., Mecham, R. P., and Ramirez, F. (1994) J. Cell Biol. 124, 855–863
5. Reinhardt, D. P., Keene, D. R., Corson, G. M., Poschl, E., Bachinger, H. P., Gambee, J. E., and Sakai, L. Y. (1996) J. Mol. Biol. 256, 104–116
6. F. Segade and C. Werneck, unpublished data.