Microfibril-Associated Glycoprotein: Characterization of the Bovine Gene and of the Recombinantly Expressed Protein†

Muhammad M. Bashir,‡ William R. Abrams,‡ Joan Rosenbloom,‡ Umberto Kucich,‡ Michelle Bacarra,‡ Ming-Da Han,‡ Patricia Brown-Augsberger,§ Robert Mecham,§ and Joel Rosenbloom*,‡

Department of Anatomy and Histology and Research Center in Oral Biology, School of Dental Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, and Department of Cell Biology, Washington University, St. Louis, Missouri 63110

Received July 22, 1993; Revised Manuscript Received October 12, 1993®

ABSTRACT: Microfibrils having a diameter of 10–12 nm, found either in association with elastin or independently, are an important component of the extracellular matrix of many tissues. Because isolation of native proteins composing the microfibrils has proven difficult, information on structure/function relationships is limited. In order to extend our understanding of the 31-kDa microfibril-associated glycoprotein (MAGP), the bovine gene has been cloned and characterized and the protein has been expressed in a eukaryotic system. The compact coding portion of the gene is contained in 4.5 kbp of genomic DNA and does not appear to share any domain motifs with other known proteins. The size, amino acid composition, and sequence of the amino terminus of the secreted recombinant protein (rMAGP) all agree with values predicted by the nucleotide sequence of the cDNA used in the expression vector. The rMAGP reacts with a monospecific antibody prepared against a defined amino acid sequence of the natural molecule and reacts specifically with recombinantly produced tropoelastin, suggesting that rMAGP will be a useful reagent with which to study its interaction with other extracellular matrix components.

Microfibrils having a diameter of 10–12 nm are found in the extracellular matrix of a variety of tissues (Low, 1962). These microfibrils have been identified as a component of elastic fibers in which they are located primarily around the periphery of an amorphous elastin core, but they are also to some extent interspersed within it (Fahrenbach et al., 1966; Greenlee et al., 1966; Karrer & Cox, 1961). While the elastic properties are attributable to the amorphous elastin, it has been postulated that the microfibrils act as an organizing scaffold for the deposition of tropoelastin molecules which are subsequently cross-linked during fibrillogenesis (Greenlee & Ross, 1967). However, in several nonelastic tissues, such as the ciliary zonule and the periodontal ligament, morphologically identical microfibrils are also found, and here they may serve an anchoring function (Cleary, 1987).

At high magnification, the microfibrils appear to have a significant substructure, suggesting that they may be composed of more than one protein. This conclusion is supported by biochemical data, but because of the difficulty of obtaining pure microfibrils and of extracting the constituent proteins the exact composition of the microfibrils remains to be defined (Greenlee & Ross, 1967). When bovine nuchal ligament, a tissue rich in microfibrils, was extracted with reductive saline, five major bands having molecular masses of 340, 78, 70, 31, and 25 kDa were identified upon gel electrophoresis (Gibson et al., 1989). Subsequent molecular cloning studies have indicated that the 340-kDa protein corresponds to the 350-kDa protein, named fibrillin, identified in culture media of human fibroblasts (Sakai et al., 1986). Two different human fibrillin genes have been identified. One located on chromosome 15q15-21 has been linked to the Marfan syndrome (Lee et al., 1991; Maslen et al., 1991), and the other on chromosome 5q23-31 has been linked to congenital contractural arachnodactyly (Lee et al., 1991). The available partial sequence data reveals that the fibrillins have a modular primary structure in which blocks of EGF repeats alternate with regions homologous to sequences in TGF-β binding protein (Lee et al., 1991; Maslen et al., 1991). The 31-kDa protein, termed microfibril-associated glycoprotein (MAGP),[1] has also recently been cloned (Gibson et al., 1991). The deduced amino acid sequence indicates that MAGP contains two structurally distinct regions: the amino-terminal half of the protein is rich in glutamine, proline, and acidic amino acids whereas the carboxy-terminal half contains all 13 of the cysteine residues and most of the basic amino acids. MAGP can be extracted in relatively small amounts from tissues by several solvents, provided they contain a strong reducing agent, suggesting that intermolecular disulfide bonding is an important feature of the association of the polypeptide chains in the fibrils.

In order to further our understanding of the gene structure of proteins composing the microfibrils and to provide a source of the protein for further detailed studies, we have cloned and characterized the bovine MAGP gene and have purified and characterized the recombinant protein expressed in a eukaryotic system.

EXPERIMENTAL PROCEDURES

Genomic Library Screening and Clone Characterization. A bovine genomic library (Clontech) was screened with the 936 bp cDNA clone cM32 encoding MAGP (Gibson et al., 1991). Two clones, which were subsequently shown to

† Supported by National Institutes of Health Grants AR41474, AR20553, HL26499, and HL41926.

* Corresponding author, at the Department of Anatomy and Histology, School of Dental Medicine, University of Pennsylvania, Philadelphia, PA 19104. Phone: 215-898-8994; FAX: 215-573-2324.
‡ University of Pennsylvania.
§ Washington University.

® Abstract published in *Advance ACS Abstracts*, December 15, 1993.

[1] Abbreviations: MAGP, microfibril-associated glycoprotein; rMAGP, recombinant microfibril-associated glycoprotein; PTC, phenylthiocarbamyl; PITC, phenyl isothiocyanate; RSA, rabbit serum albumin; PBS, phosphate-buffered saline; PGAP, pyroglutamate aminopeptidase.

0006-2960/94/0433-593$04.50/0

encompass the entire coding portion of the gene, were identified and purified to homogeneity by successive rounds of plaque purification. The clones were characterized by restriction endonuclease mapping, and restriction fragments were isolated after electrophoresis on 1% agarose gels and subcloned into pUC19. The inserts were partially sequenced by the Sanger dideoxynucleotide chain-termination method using an ABI 373A automated DNA sequencer in order to compare the genomic sequence with that of the cDNA and to determine the exon/intron structure of the gene (Sanger et al., 1977). Oligonucleotide primers were synthesized on a Milligen Model 7500 DNA synthesizer (Burlington, MA) and purified by high-pressure liquid chromatography.

Primer Extension and S1 Mapping. Poly(A^+) RNA was isolated from fetal bovine ligamentum nuchae and fetal aorta by the guanidine isothiocyanate method followed by oligo-(dT) affinity chromatography (Chirgwin et al., 1979). A synthetic antisense oligonucleotide, 5′-CAGGCAGGAA-GAGCAGGAAGAGGGA- 3′, corresponding to bases 13–37 in Figure 2, after 5′-labeling with ^{32}P, was hybridized to 0.5 μg of poly(A^+) RNA. The primer was extended using reverse transcriptase for 45 min at 42 °C, and the reaction products were ethanol-precipitated and dissolved in 0.1 M NaOH/1 mM EDTA (Chretien et al., 1988).

S1 mapping was performed according to published protocols with slight modifications (Lompre et al., 1984). The single-stranded 5′-labeled probe for S1 analyses was prepared as follows. The same antisense oligonucleotide used in primer extension was hybridized to alkali-denatured double-stranded genomic DNA and extended by T7 DNA polymerase. The extended product was digested with restriction endonuclease *Pvu*I, corresponding to a site present 294 bp 5′ of the oligonucleotide primer. The restriction fragments were separated by alkaline low-melt agarose gel electrophoresis, and the 294-base-long single-stranded probe was identified by autoradiography, excised, and eluted from the gel. This probe was hybridized to the RNA at 52 °C for 16 h in 75% formamide/0.4 M NaCl. At the end of the hybridization, the reaction mixtures were immediately diluted 10-fold with ice-cold S1 buffer containing 300 units/mL S1 nuclease. Digestion was then carried out for 1 h at 37 °C. The reaction products were extracted with phenol/chloroform and precipitated with ethanol.

The sizes of the reaction products from the primer extension and S1 nuclease digestions were determined by electrophoresis on DNA sequencing gels.

Southern Analysis. High molecular weight bovine DNA was digested with restriction endonucleases, and samples were subjected to electrophoresis on 0.8% agarose gels. After transfer of the fragments to nitrocellulose filters, the blots were hybridized to radiolabeled full-length bovine MAGP cDNA clone cM32 (Gibson et al., 1991; Southern, 1975).

Construction of the Mammalian Expression Vector. The full-length MAGP cDNA contained in pBluescript was digested with *Eco*RI/*Sma*I, which releases an insert consisting of a 25 bp 5′-untranslated segment, the entire 552 bp coding sequence, and a 28 bp 3′-untranslated segment. The entire fragment was cloned into the mammalian expression vector pCMV5 (Figure 5), which uses the promoter/enhancer region of the major immediate early gene of the human cytomegalovirus to drive expression, the transcription termination and polyadenylation region of the human growth hormone gene, and the SV40 virus DNA replication origin and early region enhancer from plasmid pcD-X (Anderson et al., 1989). HeLa S3 cells (American Type Culture CCL2.2) cultured at 37 °C in Hams F12 supplemented with 10% fetal calf serum were plated at a density of 1×10^6 cells per 100-mm dish 24 h before transfection. The recombinant expression plasmid and the selection plasmid (pRSV-neo) at molar ratios of 10:1 or 20:1 were coprecipitated with $CaPO_4$, and 10 or 20 μg, respectively, was applied to the cells (Gorman, 1985). After 48 h, the cells were trypsinized and plated 1:30 using selective medium containing 750 μg/mL gentamycin (G418). Neomycin-resistant colonies, visible after about 3 weeks, were cloned, grown up in quantity, and tested for MAGP expression.

Amino Acid Analysis. Samples of HPLC-purified bovine recombinant MAGP were transferred to two 6 × 50 mm acid-washed hydrolysis tubes, and the samples were dried under vacumn. One tube was treated with freshly prepared performic acid to oxidize the cysteine to cysteic acid prior to vapor-phase hydrolysis performed in 6 N HCl containing 1% phenol as a scavenger (v/v) at 110 °C for 20–24 h. The samples were dried under vacuum, neutralized with triethylamine, redried, and derivatized at room temperature with phenyl isothiocyanate (PITC). The resulting phenylthiocarbamyl (PTC)-amino acids were separated by reverse-phase high-pressure liquid chromatography (RPHPLC; Waters PicoTag column, 3.9 × 150 mm) and monitored by UV absorbance at 254 nm. The PTC-amino acids in the hydrolysates were identified and quantified by comparing their retention times and peak areas, respective to those obtained for a freshly derivatized standard mixture of amino acids (Bidinmeyer et al., 1984).

N-Terminal Sequence Analysis. MAGP was reduced with 0.01% tributylphosphine, electrophoresed, and electroblotted to PVDF-P (Millipore, Milford, MA). The protein was stained with Sulfarhodamine B (Pappin et al., 1990) and the single band of bovine recombinant MAGP on PVDF-P cut out with a scalpel. The predicted amino terminus of MAGP is glutamine, and it was initially observed to be blocked. Glutamine readily rearranges in the presence of weak acid to form pyroglutamic acid, which is not subject to Edman degradation (Bodansky, 1988). Pyroglutamate aminopeptidase (Boeringer Mannheim) was used to remove the N-terminal pyroglutamate *in situ* (Andrews et al., 1991). PVDF-P containing the electroblotted MAGP sample was moistened with ~25 μL of 50% methanol, incubated for 30 min at 37 °C with 1.0 mL of 0.5% PVP40 in 0.1 M acetic acid, and washed with 10 × 1.0 mL of water. The membrane was then incubated with 1.0 mg/mL PGAP in 0.5 mL of reaction buffer [50 mM phosphate, pH 8.0, 5 mM DTT, 10 mM EDTA (dicalcium, disodium salt), 5% glycerol] for 12 h at 4 °C, and at room temperature for 3 h. After washing with 10 × 1.0 mL of water, the membrane was dried under vacuum, sparged with argon, and stored at −20 °C until sequence analysis using the adsorptive protocol on a Millipore ProSequencer 6625 (Long et al., 1992). PTH-amino acids were separated on-line using a 3.9 × 300 mm SequeTag column with limit buffers of 60 mM ammonium acetate, pH 4.71, and 100% acetonitrile and identified by comparison to 50 pmol of PTH standards (Sigma). After PGAP treatment, recombinant bovine MAGP yielded to the Edman chemistry.

Immunologic Analysis of rMAGP. The synthetic peptide CVIPAPTLEPGTV (residues 74–85 of bovine MAGP) was prepared using T-Boc chemistry on an ABI 431A peptide synthesizer, and the sequence was confirmed by sequence analysis using an ABI 473 protein sequencer. For antibody production, the synthetic peptide was coupled to rabbit serum albumin (RSA) using MBS (*m*-maleimidobenzoic acid *N*-hydroxysuccinimide ester) at a substitution of 7.5 mg of peptide/mg of RSA. One milligram of the peptide–RSA conjugate in 1 mL of Freund's complete adjuvant was injected

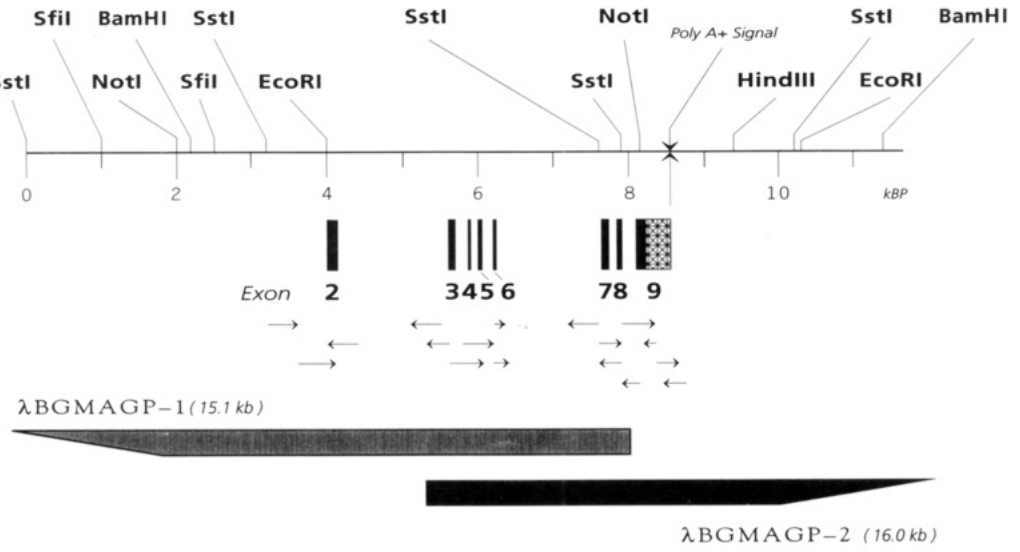

FIGURE 1: Diagram of bovine MAGP gene including positions of genomic clones and sequencing strategy. Genomic clones BGMAGP1 and BGMAGP2, isolated as described in the Experimental Procedures, were subjected to restriction enzyme analysis and sequenced by the dideoxy method using synthetic oligonucleotides (→). The untranslated portion of exon 9 is cross-hatched.

subcutaneously at 10 different sites along the backs of rabbits. Beginning at 3 weeks after initial immunization, the rabbits were given biweekly booster injections of 1 mg of peptide–RSA in 100 μL of Freund's incomplete adjuvant. Immune serum from each rabbit was diluted with caprylic acid (0.7 mL of caprylic acid/mL of serum), stirred for 30 min, and centrifuged at 5000g for 10 min. The supernatant was decanted and dialyzed against two changes of phosphate-buffered saline (PBS) overnight at 4 °C. The antibody solution was then affinity purifed by passing it over a column containing the peptide coupled to Affi-Gel 10 affinity support. Bound antibody was eluted with 0.2 M glycine, pH 2.3, immediately dialyzed against PBS, and concentrated to 1 mg/mL prior to storage at –70 °C.

This antibody was then tested for reactivity with rMAGP by Western immunoblotting and ELISA as described in the legends to Figures 6 and 8.

RESULTS

Structure of the Gene. The two genomic clones, encompassing 28 kbp, were characterized by restriction endonuclease mapping and comparison of the DNA sequence to the cDNA (Figure 1). These initial analyses demonstrated that the entire published cDNA sequence was contained within the cloned genomic DNA and permitted the definition of the exon/intron structure of the coding and 3′-untranslated portions of the gene. This region of the gene is composed of 8 exons (Figures 1 and 2). There was 100% agreement between the sequence contained in the designated exons and the corresponding cDNA, including the 3′-untranslated segment. Exons 2–8 range in size from 27 to 144 bp. Exon 2 consists of a 5′-untranslated segment and a segment encoding a portion of the signal sequence, the remainder of which is encoded in exon 3. Exon 9 encodes the carboxy end of the protein and the 3′-untranslated segment. Exon/intron borders (not shown) conform to the consensus sequences deduced from analysis of a large number of eukaryotic genes (Breathnach & Chambon, 1981). Exon/intron borders divide codons, usually between the first and second bases.

Southern Analyses. In order to verify the legitimacy of the cloned genomic DNA, Southern analyses of bovine genomic DNA were carried out using the cDNA clone cM32, which spans exons 2–9, as probe (Figure 3). The sizes of the observed restriction fragments observed in this analysis were completely

```
-107                gctgaattctgggagcagaagcggccgtgcggccctggggagcagct  -61
 -60  cgcccctcctctttctccagctgtccccagacacctgcccagcgtgctcactgagcagtc   -1

   1  ATG AGA GCT GCC TCC CTC TTC CTG CTC TTC CTG CCT GCA GGC CTG     45
   1   M   R   A   A   S   L   F   L   L   F   L   P   A   G   L    15

  46  CTG GCT CAG GGC CAG TAT GAT CTG GAC CCC CTG CCT CCG TAC CCA     90
  16   L   A   Q   G   Q   Y   D   L   D   P   L   P   P   Y   P    30

  91  GAC CAC GTG CAG TAT ACC CAC TAC AGC GAG CAG ATC GAG AAT CCG    135
  31   D   H   V   Q   Y   T   H   Y   S   E   Q   I   E   N   P    45

 136  GAC TAC TAT GAC TAC CCA GAG ATG ACC CCT CGG CCG CCC GAG GAG    180
  46   D   Y   Y   D   Y   P   E   M   T   P   R   P   P   E   E    60

 181  CAG TTC CAG TTC CAG TCC CAG CAG CAA GTC CAG CAG GAA GTC ATC    225
  61   Q   F   Q   F   Q   S   Q   Q   Q   V   Q   Q   E   V   I    75

 226  CCT GCC CCC ACC TTA GAA CCG GGG ACT GTG GAG ACG GAG CCC ACG    270
  76   P   A   P   T   L   E   P   G   T   V   E   T   E   P   T    90

 271  GAG CCG GGA CCT CTG GAC TGC CGC GAG GAG CAG TAC CCG TGC ACC    315
  91   E   P   G   P   L   D   C   R   E   E   Q   Y   P   C   T   105

 316  CGC CTC TAC TCC ATA CAC AAG CCC TGC AAG CAG TGT CTC AAC GAG    360
 106   R   L   Y   S   I   H   K   P   C   K   Q   C   L   N   E   120

 361  GTC TGC TTC TAC AGC CTC CGC CGC GTC TAC GTC GTC AAT AAG GAG    405
 121   V   C   F   Y   S   L   R   R   V   Y   V   V   N   K   E   135

 406  ATC TGC GTC CGG ACG GTC TGC GCC CAG GAG GAG CTC CTG CGG GCT    450
 136   I   C   V   R   T   V   C   A   Q   E   E   L   L   R   A   150

 451  GAC CTG TGC CGT GAC AAG TTC TCC AAG TGC GGC GTG CTG GCC AGC    495
 151   D   L   C   R   D   K   F   S   K   C   G   V   L   A   S   165

 496  AGT GGC CTG TGC CAG TCT GTC GCG GCC GCC TGT GCC AGG AGC TGT    540
 166   S   G   L   C   Q   S   V   A   A   A   C   A   R   S   C   180

 541  GGG GGC TGC TAG                                            552
 181   G   G   C  Stop                                           183

ggggagctggccccccgccagcccggggcccccaggctctgcctgacctggtgctttccc
cccgtcccacgttcctctcgtcttgttagagggctgcgggcctggggccaccccggagc
ctcgaggggctgccgaagggctagtctctgcctccgggcgggggacgacccgggtgctc
tgtgggacctgggcccctgccatgccttccttgtccccctccaggacctgtcccaccccc
gaggtaggctgtgacccccaccccagctggtctgcttgaatctcctacagccccctgggcg
gagaccacctttgttttatacaaaattaaaaacaggtttttacg
```

FIGURE 2: Nucleotide sequence and translated amino acid sequence of bovine MAGP. Division of the sequence into exons is indicated. An antisense oligonucleotide corresponding to the underlined sequence was used in the primer-extension experiments illustrated in Figure 4. Amino acid residue 18 (Q), which is underlined, is the amino terminus of the mature protein. A potential polyadenylation signal, **aattaaa**, is in boldface.

consistent with those predicted by the cloned genomic DNA (Figure 1), thus substantiating the validity of the genomic clones. The simplicity of the hybridization pattern also supports the conclusion that MAGP is found as a single-copy gene in the bovine genome.

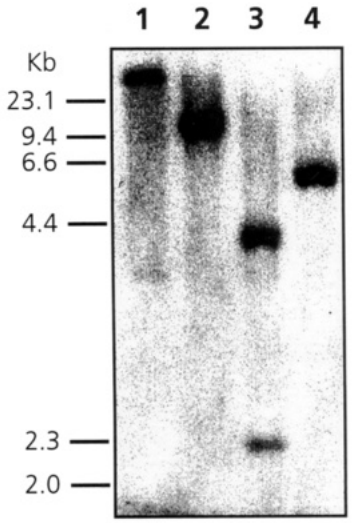

FIGURE 3: Southern blot analysis of bovine genomic DNA. Southern blot analyses were carried out as described in the Experimental Procedures using the cDNA clone cM32 encompassing exons 2–9, and the resulting autograph was digitized. Observed restriction fragments are consistent with the map of the cloned DNA given in Figure 1 (a 300 bp *Sst*1 fragment ran off the gel). Lane 1, *Hin*dIII; lane 2, *Bam*HI; lane 3, *Sst*1; lane 4, *Eco*RI.

Primer Extension and S1 Analyses. In order to identify transcription initiation site(s) and to locate the potential promoter region, primer extension and S1 protection analyses were carried out. These experiments used an antisense oligonucleotide corresponding to translated nucleotides 13–37 (Figure 2) in the primer extension experiments and a probe extending 300 nucleotides 5′ of this sequence in the S1 protection experiments (Figure 4). Electrophoretic analysis of the S1 digestion products demonstrated that a single product of about 150 bases was produced. However, when the primer-extended products were analyzed in a similar fashion, multiple products larger than this were seen. These experiments suggested that additional transcribed sequences existed in one or more exons located 5′ of the genomic segment used to generate the S1 probe, but the locations of the additional sequences in the genomic DNA and of the provisional promoter remain to be determined. The sequence found to be protected in the S1 assays was determined by sequencing the corresponding segment of genomic DNA and a provisional exon/intron border identified. This additional segment has been included as part of the 5′-untranslated segment, where it appears as residues –26 to –107 in Figure 2.

Production of rMAGP. Although tropoelastin has been produced in quantity in a prokaryotic system (Indik et al., 1990), it had to be synthesized as a fusion protein since the free polypeptide was rapidly degraded. Fortuitously, tropoelastin contains no internal methionine residues, so the tropoelastin could be prepared rather readily by column chromatographic separation of CNBr cleavage products. However, such an approach is not available for the production of rMAGP, since it contains an internal methionine. Furthermore, MAGP contains 13 cysteines, some of which are likely to be involved in intramolecular disulfide bonds which may not form correctly in prokaryotic expression systems. Therefore, we turned to a eukaryotic expression vector, pCMV5, which is based upon the powerful cytomegalovirus promoter. The MAGP coding sequence including the signal sequence was inserted into the multiple cloning site of the vector as described in the Experimental Procedures and as illustrated in Figure 5. Permanently transfected cells were selected by neomycin resistance, and the cloned cells were tested for MAGP expression. Cells grown to near confluency

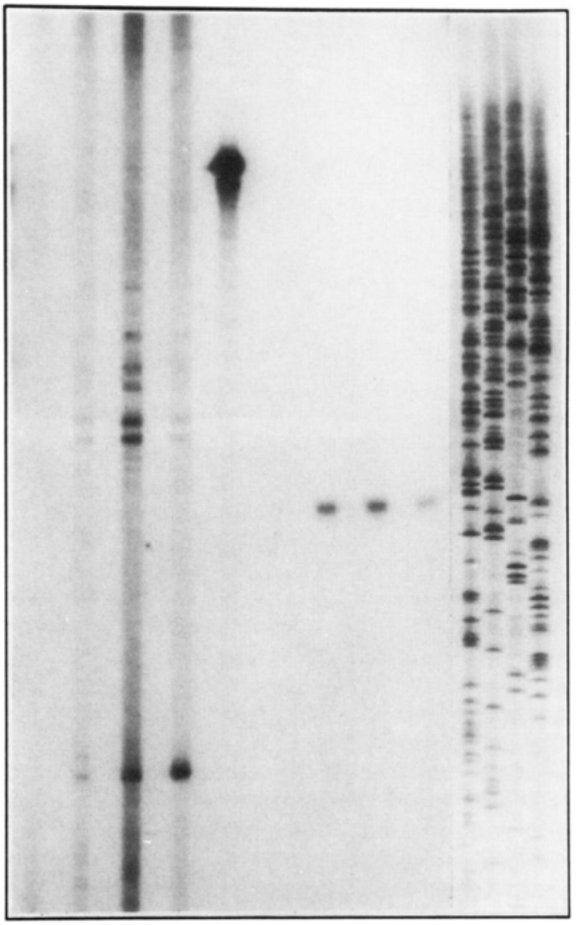

FIGURE 4: Primer extension and S1 protection analyses. Details are described in the Experimental Procedures. Primer extension analysis used an antisense oligonucleotide corresponding to translated nucleotides 13–37 (Figure 2) and an S1 probe extending 300 nucleotides 5′ of this sequence. Lanes 1–4, primer extension; lanes 5–9, S1 protection. (1) tRNA, (2) 180-day ligamentum nuchae poly(A^+) RNA, (3) 160-day ligamentum nuchae poly(A^+) RNA, (4) 150-day aorta poly(A^+) RNA, (5) undigested probe, (6) tRNA, (7) 180-day ligamentum nuchae poly(A^+) RNA, (8) 160-day ligamentum nuchae poly(A^+) RNA, (9) 150-day aorta poly$A^{(+)}$ RNA. The dideoxynucleotide sequencing reaction in lanes 10–13 was carried out with the same oligonucleotide used for primer extension, and the resulting oligonucleotides served as size markers. (10) G reaction, (11) A reaction, (12) T reaction, (13) C reaction.

in 35-mm dishes in medium containing 10% fetal calf serum were then washed in serum-free medium and incubated in serum-free medium containing 3 μCi/mL [^{3}H]proline for 24 h. After addition of bovine serum albumin to a final concentration of 10 μg/mL as a carrier protein, the proteins secreted into the media were precipitated by addition of TCA to a final concentration of 10% and subjected to SDS–PAGE and autoradiographic analysis. Of the 24 neomycin-resistant clones, 5 produced a 31-kDA protein which was provisionally identified as MAGP and which was not produced by the parent HeLa cells (representative results are shown in Figure 6A). On the basis of these incorporation experiments, two clones were selected for analysis in more detail. Near-confluent cultures were incubated for 48 h in serum-free medium, and the secreted proteins were precipitated as described above. Western immunoblotting demonstrated that the clones synthesized and secreted a 31-kDa protein which was immunoreactive with the monospecific, affinity-purified antibody generated against a peptide sequence found in MAGP (Figure 6B). As determined by Western blotting, there appeared to be relatively little difference between the clones in the rate at

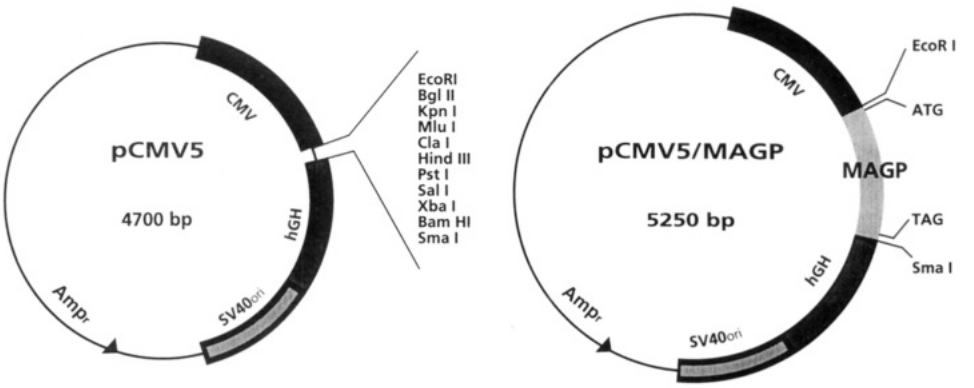

FIGURE 5: Construction of eukaryotic expression vector, pCMV5/MAGP. Plasmid pCMV5 represents the starting expression vector and contains the immediate early promoter region of the human cytomegalovirus (CMV), a polylinker containing unique sites for the indicated restriction enzymes, transcription termination and polyadenylation signals from the human growth hormone gene (hGH), and the SV40 origin of DNA replication and early region enhancer sequences ($SV40_{ori}$). This plasmid also contains an *Escherichia coli* gene encoding ampicillin resistance (Amp^r). Plasmid pCMV5/MAGP was constructed by cloning a 605 bp insert containing the entire coding sequence and portions of the 5′- and 3′-untranslated segments into the *Eco*RI and *Sma*I sites of pCMV5. The positions of the initiator methionine (ATG) and translation termination (TAG) codons are indicated.

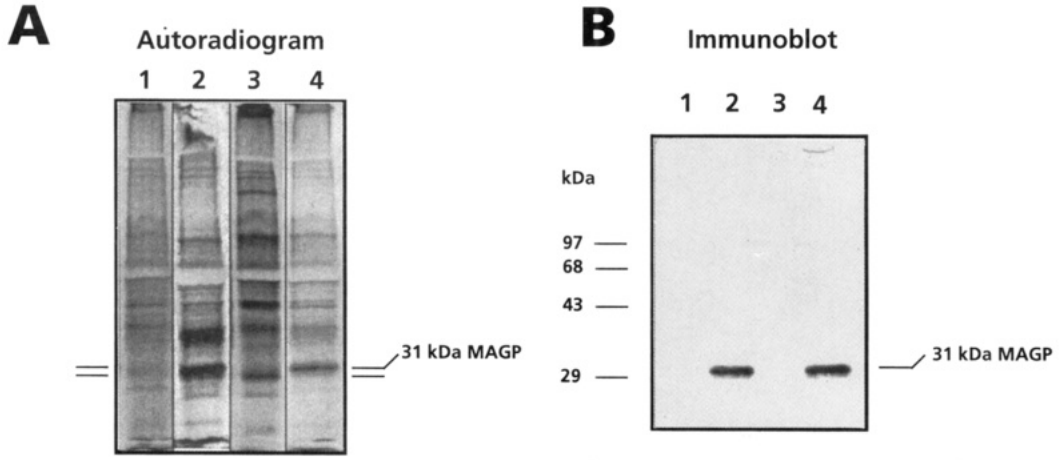

FIGURE 6: SDS–PAGE of labeled proteins and Western immunoblot of proteins secreted by neomycin-resistant clones. (A) Secreted proteins labeled with [^{3}H]proline were subjected to electrophoresis on 10% SDS–PAGE under reducing conditions followed by autoradiographic analysis. Samples from representative clones are illustrated. Lane 1, control HeLa cells; lane 2, clone 20-2-1: lane 3, clone 10-3-1; lane 4, clone 10-5-1. (B) Unlabeled secreted proteins were subjected to electrophoresis as in (A) and then electroblotted onto nitrocellulose filters. The transferred proteins were reacted with affinity-purified antibody directed against a 12 amino acid synthetic peptide corresponding to residues 74–85 of MAGP. Lanes are the same as in (A). Two of the neomycin-resistant clones are seen to be producing MAGP.

which they synthesized MAGP, and clone 10-5-1 was arbitrarily taken for large-scale production.

Purification and Characterization of rMAGP. Cells were grown in monolayers in 150-cm^2 flasks and subjected to convenient weekly cycles of alternate incubation in serum-free medium for 3 days and serum-containing medium for 4 days. It was found that this alternation could be repeated 4 times without loss of MAGP production. Preliminary attempts to purify the MAGP revealed that it was easily lost, probably by nonspecific adherence to surfaces. To prevent such loss, chromatographic-grade urea was added to the serum-free incubation medium to a final concentration of 5 M and the proteins were concentrated 30-fold using a Amicon PM10 membrane. The MAGP was purified to homogeneity by sequential chromatography, first on a weak anion-exchange column (DEAE-8-HR, Waters) using a Pharmacia fast-flow apparatus, and then on a Shodex DE-613 hydrophilic polymethacrylate column using a Waters high-pressure liquid chromatograph. The specifics of the chromatographic procedures are given in the legend to Figure 7. Purification of the MAGP was monitored by dot-blot immunologic analysis of the fractions. The resultant preparation appeared to be homogeneous when analyzed by silver staining of SDS–PAGE gels and gave a single product on Western analysis (Figure 7).

The purified protein was characterized by amino acid analysis, and the composition (Table 1) compares favorably with that predicted by the nucleotide sequence. The deduced amino acid sequence indicates that MAGP contains a secretory signal sequence of 17 predominantly hydrophobic amino acids followed by Gln-Gly-Gln. It is possible that cleavage of the signal sequence occurs at either glutamine 18 or 20 since both positions follow the "(−3, −1)" rule for cleavage site location (Von Heijne, 1986). The cyclization of glutamine to pyrollidonecarboxylic acid prevents direct amino-terminal sequencing of the polypeptide, and preliminary experiments indicated such a blockage. In order to circumvent this difficulty and to determine the actual amino-terminal sequence of the secreted protein, the rMAGP was incubated with pyroglutamate aminopeptidase to remove the N-terminal pyroglutamate (Andrews et al., 1991). This was successfully accomplished, and the determined amino acid sequence unambiguously established that the amino terminus of the secreted MAGP corresponded to glutamine 18 (Table 1).

Enzyme-Linked Immunoabsorbent Assay (ELISA) of MAGP. The monospecific antibody preparation was used in a direct ELISA to demonstrate the potential utility of rMAGP as a standard reagent in immunologic assays. As shown in a representative ELISA curve (Figure 8), the quantitation of

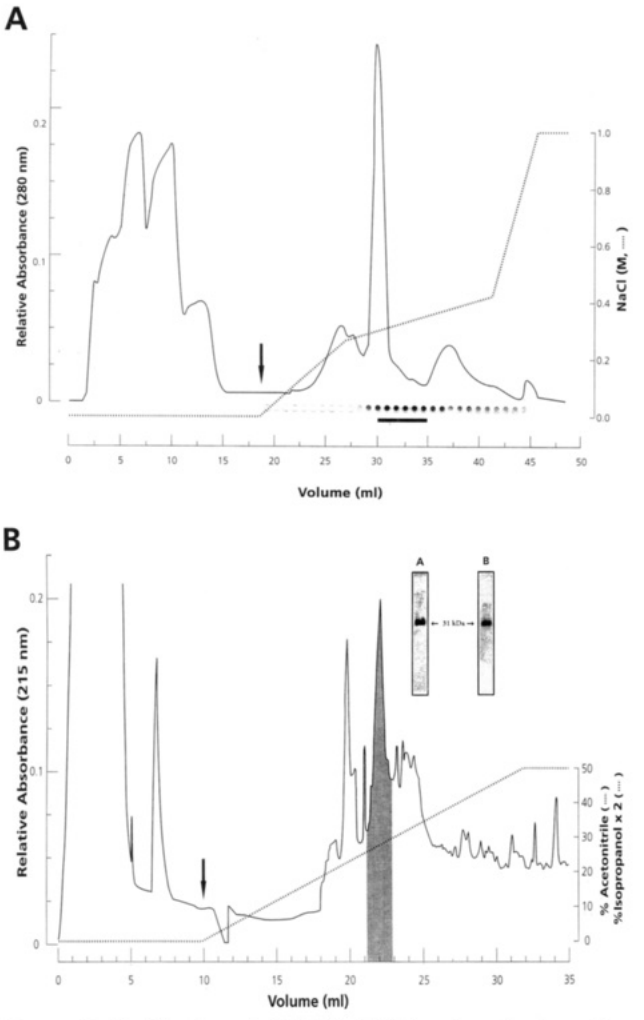

FIGURE 7: Purification of rMAGP. (A) Fractionation by anion-exchange chromatography. Medium containing MAGP secreted by permanently transfected HeLa S3 cells was clarified by centrifugation at 5000*g* for 15 min, and the supernatant was made 5 M in urea and concentrated 30-fold by ultrafiltration on a 47-mm PM10 (Amicon) membrane. The concentrated supernatant was applied to a Protein Pak DEAE-8-HR (Waters) column equilibrated in 25 mM triethylamine, pH 7.0/5.0 M urea connected to an FPLC chromatograph (Pharmacia) using successive loadings with a 10-mL superloop. The column was washed with starting buffer (0.5 mL/min) until no further absorbance eluted, and then the gradient was initiated (arrow). Fractions of 1.5 mL were collected. The dashed line indicates the molar concentraton of NaCl used to elute the MAGP. Absorbance at 280 nm is represented by the solid line. Each fraction was analyzed by spotting 200 μL, diluted with 300 μL of H_2O, onto nitrocellulose and probing for rMAGP with polyclonal, monospecific antibody. The resulting dot blot was digitized and is shown as an inset positioned over the volume in which it was collected. The active fractions (dark bar under the immuno-dot blot) were pooled. The chromatographic tailing of MAGP is probably due to glycosylation and possibly other posttranslational modifications (*6*, *7*). (B) Fractionation of rMAGP with reverse-phase HPLC. Pooled immunologically active fractions from anion-exchange chromatography were loaded directly onto a RSpak Shodex DE-613 (6 mm × 150 mm) hydrophilic polymethacrylate column with no additional functional group using a Waters 625 pump and monitored at 215 nm. The column was equilibrated with 0.01 N HCl at 40 °C and a flow rate of 0.5 mL/min. The column was washed until urea and salts were removed from the sample and the absorbance had returned to base-line levels. Elution was effected by a binary linear gradient (arrow indicates gradient initiation), which increased both acetonitrile and 2-propanol concentration simultaneously at a rate of 1.11%/min and 0.56%/min, respectively, for 45 min at a constant concentration of 0.01 M HCl. One milliliter fractions were collected. The inset shows the silver-stained and Western blot of the gray area, indicating a single band of the anticipated size (31 kDa).

MAGP is effective over a concentration range of 10–100 ng/mL.

Table 1: Analytical Analysis of Recombinant Bovine MAGP

amino acid analysis			amino-terminal sequence analysis			
amino acid	observed residues/ molecule	expected residues/ molecule	residue no.	observed amino acid	yield (pmol)	expected amino acid
Cya[a]	7	13	1	pG[b]		Q
D	12	11	2	G[c]	14.84	G
E	28	32	3	Q	1.68	Q
S	10	9	4	Y	3.7	Y
G	12	7	5	D	11.14	D
H	3	3	6	L	0.91	L
R	10	9	7	D	7.62	D
T	9	8	8	P	1.31	P
A	10	8	9	L	2.05	L
P	21	17	10	P	1.56	P
Y	7	11	11	P		P
V	11	12	12	Y	2.61	Y
M	1	1	13	X		P
I	7	4	14	D	1.83	D
L	13	12	15	X		H
F	5	4	16	V	0.75	V
K	6	5				

[a] Cya = cysteic acid. Cysteine content determined as cysteic acid after performic acid oxidation. [b] pG = pyroglutamate. As isolated, rbMAGP did not yield to Edman Degradation, suggesting a blocked amino terminus. This residue was predicted to be pyroglutamate due to exposure to acid during purification. Pyroglutamate aminopeptidase treatment removed this residue. [c] This is the penultimate residue since sample was subjected to pyroglutamate aminopeptidase treatment.

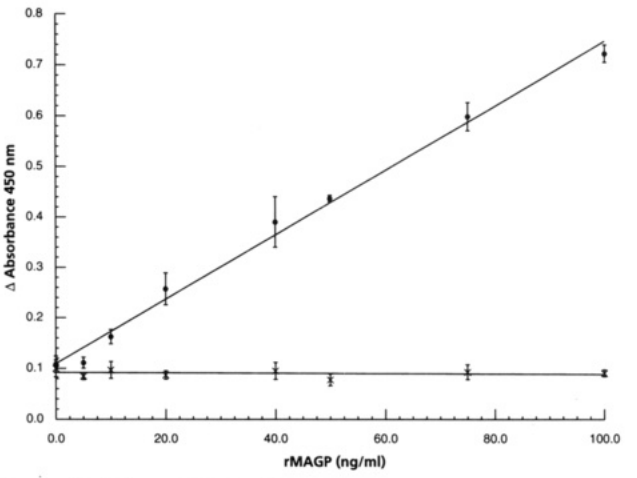

FIGURE 8: Direct ELISA of rMAGP. Varying concentrations of bovine rMAGP were absorbed to the wells of polystyrene microtiter plates by overnight incubation in 0.1 M carbonate, pH 9.6, at 4 °C. After blocking with 5% Blotto, the bound MAGP was detected by reactivity with monospecific, polyclonal, affinity-purified rabbit antibody to the peptide VIPAPTLEPGTV (residues 74–85 of bovine MAGP) followed by horseradish peroxidase complex-coupled goat anti-rabbit IgG (●). Normal rabbit serum control (×). Determinations were done in triplicate, and the averages ± 1 standard deviation are plotted.

Binding of rMAGP to Recombinant Tropoelastin. The potential interactions of tropoelastin and elastin with microfibrillar components are of particular interest since such interactions may be of critical importance in the assembly of functional elastic fibers. In order to determine whether MAGP and tropoelastin bind to one another, a modified ELISA was performed. Wells were coated with human recombinant tropoelastin prepared as previously described (Indik et al., 1990), nonspecific binding was blocked with Blotto, and the coated wells then incubated with MAGP, followed by sequential addition of anti-MAGP antibody and horseradish peroxidase-coupled goat anti-rabbit IgG (Figure 9). The results of these experiments clearly demonstrated that specific binding between MAGP and tropoelastin takes place, although

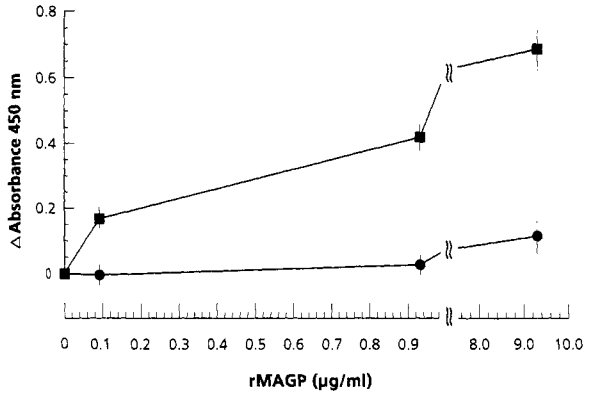

FIGURE 9: Binding of rMAGP to recombinant tropoelastin. Recombinant tropoelastin (0.5 μg/mL in 0.1 M carbonate, pH 9.6) was absorbed to the wells of polystyrene microtiter plates by incubation overnight at 4 °C. After blocking with 5% Blotto, rMAGP was added at the indicated concentrations and incubated for 1 h at 37 °C. The specifically bound rMAGP was then detected as described in Figure 8 (■). Control wells were not coated with tropoelastin (●).

more extensive studies are clearly needed to characterize this interaction in detail.

DISCUSSION

Although cDNA encoding several microfibrillar proteins have been cloned (Lee et al., 1991; Maslen et al., 1991; Gibson et al., 1991), there is only fragmentary knowledge concerning the structure of the corresponding genes. The fibrillins compose a family in which the member genes appear to have been assembled predominantly from two motifs, one homologous to a repeat found in EGF (Carpenter & Cohen, 1990) and a second homologous to a segment in TGF-β binding protein (Kanzaki et al., 1990). Together these motifs acount for greater than 90% of the known fibrillin sequences. The exons encoding these individual homologous sequences are also organized in the fibrillin genes in a manner similar to that found in the EGF and TGF-β binding protein genes. No information was previously available with respect to the gene structure of other putative microfibrillar proteins which prompted the present studies. Molecular mechanics modeling of MAGP, which is a comparatively small protein, using the QUANTA/CHARMm parameters of the Polygen Corp. (Momany & Rone, 1992; our results not shown), suggested that the protein may be organized into three domains (minus the signal sequence): (1) an amino-terminal globular segment (residues 18–49), (2) an extended proline-rich segment (residues 50–95), and (3) a relatively large, globular domain which may be heavily disulfide bonded (residues 96–183). This analysis, in turn, suggested that the potential domain structure may be a reflection of the gene structure. However, determination of the MAGP gene structure (Figures 1 and 2) failed to support this rather simple hypothesis. Nevertheless, some elements of the general idea may hold as follows: the first globular domain of the protein can be considered to be encoded by exons 3 and 4, the second extended domain to be encoded by exons 5 and 6, and the comparatively large carboxy domain to be encoded by exons 7–9. It is interesting to note that all of the 13 cysteine residues in the molecule are found in the carboxy globular domain, with exon 7 containing five cysteines, exon 8 containing two cysteines, and exon 9 containing six cysteines. This analysis may provide some clue to potential disulfide bond pairing, which will be difficult to determine because of the comparatively large number of cysteines. We hypothesize that intrachain disulfide bonds will occur between residues encoded in the same exon.

When primer extension and S1 nuclease protection experiments were carried out, the results indicated the presence of one or more noncoding exons in the 5′ portion of the transcript, since the primer extension products extended well beyond the 5′ end of the S1-protected product. Multiple primer extension products were obtained (Figure 4, lanes 2–4), suggesting that transcription may be initiated at several sites. However, this region is extremely GC-rich and some of these products may be the result of nonspecific termination, and thus, further experiments are necessary to determine the nature of these products. The S1-protected product located an exon border which conformed to the consensus sequence, and we have provisionally numbered the exons starting with 2, assuming that only one more 5′ exon will be found, although this might not hold true. Experiments are in progress to locate and characterize the remainder of the upstream transcribed sequence, but the relative ease with which it can be identified is uncertain since first introns are of variable size and may be quite large.

Characterization and determination of the structure of microfibrillar proteins themselves has been limited by their availability. Isolation from tissues can only be achieved through the use of strong reducing agents usually in the presence of a strong chaotropic agent, and the proteins are obtained in denatured form. Nontransfected cultured cells synthesize relatively small amounts of the proteins, making them unsuitable sources for the preparation of substantial amounts of the proteins. Production by recombinant methods is an obvious potential solution to the problem. However, because many of the proteins contain intrachain disulfide bonds, synthesis in prokaryotic systems is likely to be of limited value. Therefore, production was attempted in a permanently transfected eukaryotic cell system using a vector based upon the powerful promoter of the immediate early gene of the human cytomegalovirus to drive expression. The permanently transfected HeLa S3 cells are hardy and readily grown to confluency. The confluent cultures can be incubated in alternating periods of serum-free and serum-containing medium, resulting in economies of scale and a ready supply of starting material in the serum-free medium. With this system, it was possible to prepare essentially homogeneous rMAGP as determined by gel electrophoresis and Western blotting. The identity of the protein was confirmed by amino acid analysis, and the sequence of the amino terminus of the secreted protein, which had not previously been obtained because the amino terminus is blocked, was determined. The present determination was accomplished by first enzymatically removing the terminal pyroglutamate. The results clearly demonstrated that the signal sequence had been removed during biosynthesis in the cells by cleavage between Ala-17 and Gln-18. In this system, approximately 100 μg of purified MAGP can be obtained from eight 150-cm^2 flasks of confluent cells ($\sim 2 \times 10^7$ cells/flask) incubated in serum-free medium for 3 days. The rMAGP reacts with monospecific antibody generated against a sequence in natural MAGP, and the recombinant protein should be suitable for future protein structural studies, including cysteine pairing in disulfide bond formation.

Ultrastructural and biochemical studies have suggested that the microfibrils are complex structures (Fahrenbach et al., 1966; Greenlee et al., 1966; Karrer & Cox, 1961; Greenlee & Ross, 1967; Cleary, 1987; Gibson et al., 1989), and clarification of the interactions among microfibrillar components themselves and with other matrix macromolecules including elastin is fundamental to understanding the function of the microfibrils. As demonstrated in Figure 9, rMAGP reacts specifically with recombinantly produced tropoelastin. This is an important observation since elastic fibers always

contain microfibrils, and it has been hypothesized, on purely circumstantial grounds, that these microfibrils act as a scaffold upon which the elastin is assembled. The present study, which is the first demonstration of a specific interaction between a microfibrillar protein and elastin, lends credibility to this idea, and the rMAGP should prove useful in more detailed studies of the interaction between itself and other microfibrillar proteins as well as with elastin.

REFERENCES

Anderson, S. A., Davis, D. L., Dahlback, H., Jornvall, H., & Russell, D. W. (1989) *J. Biol. Chem. 264,* 8222–8229.

Andrews, D. W., Girardi, M., & Mark, J. (1991) in *Techniques in Protein Chemistry II* (Villafranca, J. J., Ed.) pp 75–79, Academic Press, New York.

Bidinmeyer, B. A., Cohen S. A., & Tarvin, T. L. J. (1984) *Chromatography 336*, 93–104.

Bodansky, M. (1988) *Peptide Chemistry*, p 108, Springer-Verlag, New York.

Breathnach, R., & Chambon, P. (1981) *Annu. Rev. Biochem. 50*, 349–383.

Carpenter, G., & Cohen, S. (1990) *J. Biol. Chem. 265*, 7709–7712.

Chirgwin, J. M., Przybyla, A. A., McDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Chretien, S., Dubart, A., Beaupain, D., Raich, N., Grandchamp, B., Rosa, J., Goossens, M., & Roseo, P.-H. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6–10.

Cleary, E. G. (1987) in *Connective Tissue Disease. Molecular Pathology of the Extracellular Matrix* (Uitto, J., & Perejda, A. J., Eds.) pp 55–81, Marcel Dekker, New York.

Fahrenbach, W. H., Sandberg, L. B., & Cleary, E. G. (1966) *Anat. Rec. 155*, 563–576.

Gibson, M. A., Kumaratilake, J. S., & Cleary, E. G. (1989) *J. Biol. Chem. 264*, 4590–4598.

Gibson, M. A., Sandberg, L. B., Grosso, L. E., & Cleary, E. G. (1991) *J. Biol. Chem. 266*, 7596–7601.

Gorman, C. (1985) in *DNA Cloning, A Practical Approach*, Vol. IV, pp 143–190, IRL Press, Oxford.

Greenlee, T. K., & Ross, R. (1967) *J. Ultrastruct. Res. 18*, 354–376.

Greenlee, T. K., Jr., Ross, R., & Hartman, J. L. (1966) *J. Cell Biol. 39*, 59–71.

Indik, Z., Abrams, W. R., Kucich, U., Gibson, C. W., Mecham, R.P., & Rosenbloom, J. (1990) *Arch. Biochem. Biophys. 280*, 80–86.

Kanzaki, T., Olofsson, A., Moren, A., Wernstedt, C., Hellman, U., Miyazono, K., Claesson-Welsh, L., & Heldin, C.-H. (1990) *Cell 61*, 1051–1061.

Karrer, H. E., & Cox, J. (1961) *J. Ultrastruct. Res. 4*, 420–454.

Lee, B., Godfrey, M., Vitale, E., Hori, H., Mattei, M.-G., Sarfarazi, M., Tsipouras, P., Ramirez, F., & Hollister, D. W. (1991) *Nature 352*, 330–334.

Lompre, A. M., Nadal-Ginard, B., & Mahdavi, V. (1984) *J. Biol. Chem. 259*, 6437–6443.

Long, D., Wilcox, W. C., Abrams, W. R., Cohen, G. H., & Eisenberg, R. J. (1992) *J. Virol. 66*, 6668–6685.

Low, F. N. (1962) *Anat. Rec. 142*, 131–137.

Maslen, D., Corson, G. M., Maddox, B. K., Glanville, R. W., & Sakai, L. Y. (1991) *Nature 352*, 334–337.

Momany, F. A., & Rone, R. (1992) *J. Comput. Chem. 13*, 888–900.

Pappin, J. C. D., Coull, J. M., & Koster, H. (1990) *Anal. Biochem. 187*, 10–19.

Sakai, L. Y., Keene, D. R., & Engvall, E. (1986) *J. Cell Biol. 106*, 2499–2509.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Southern, E. M. (1975) *J. Mol. Biol. 98*, 503–517.

von Heijne, G. (1986) *Nucleic Acids Res. 14*, 4683–4690.