JOURNAL OF BONE AND MINERAL RESEARCH
Volume 16, Number 9, 2001
© 2001 American Society for Bone and Mineral Research

Matrix Gla Protein in *Xenopus laevis*: Molecular Cloning, Tissue Distribution, and Evolutionary Considerations*

M.L. CANCELA,[1] M.C.P. OHRESSER,[1,2] J.P. REIA,[1] C.S.B. VIEGAS,[1] M.K. WILLIAMSON,[3] and P.A. PRICE[3]

ABSTRACT

Matrix Gla protein (MGP) belongs to the family of vitamin K–dependent, Gla-containing proteins and in higher vertebrates, is found in the extracellular matrix of mineralized tissues and soft tissues. MGP synthesis is highly regulated at the transcription and posttranscription levels and is now known to be involved in the regulation of extracellular matrix calcification and maintenance of cartilage and soft tissue integrity during growth and development. However, its mode of action at the molecular level remains unknown. Because there is a large degree of conservation between amino acid sequences of shark and human MGP, the function of MGP probably has been conserved throughout evolution. Given the complexity of the mammalian system, the study of MGP in a lower vertebrate might be advantageous to relate the onset of MGP expression with specific events during development. Toward this goal, MGP was purified from *Xenopus* long bones and its N-terminal amino acid sequence was determined and used to clone the *Xenopus* MGP complementary DNA (cDNA) by a mixture of reverse-transcription (RT)– and 5′- rapid amplification of cDNA ends (RACE)–polymerase chain reaction (PCR). MGP messenger RNA (mRNA) was present in all tissues analyzed although predominantly expressed in *Xenopus* bone and heart and its presence was detected early in development at the onset of chondrocranium development and long before the appearance of the first calcified structures and metamorphosis. These results show that in this system, as in mammals, MGP may be required to delay or prevent mineralization of cartilage and soft tissues during the early stages of development and indicate that *Xenopus* is an adequate model organism to further study MGP function during growth and development. (J Bone Miner Res 2001;16:1611–1621)

Key words: matrix Gla protein, *Xenopus laevis*, cDNA, evolution, development

INTRODUCTION

MATRIX GLA protein (MGP) is a small protein (10 kDa molecular weight) that contains five residues of the vitamin K–dependent calcium binding amino acid γ-carboxyglutamic acid (Gla). In mammals, the transcription of the MGP gene occurs in all tissues analyzed[1–3] and it is expressed at comparable levels in fetal and adult tissue,[4] suggesting that it may play a role during vertebrate development.

The expression of the MGP gene has been shown to be affected by steroid hormones such as retinoic acid[2,5,6] and the hormonally active form of vitamin D_3, 1,25-dihydroxyvitamin D_3,[2,7] as well as by growth factors and by cell proliferation.[8,9] MGP is also subject to complex

*The full length xMGP cDNA sequence was submitted to the EMBL database with the accession number AF055588.1.

[1]Centro de Ciências do Mar, Campus de Gambelas, Universidade do Algarve, Faro, Portugal.
[2]Present address: Laboratoire de Neurobiologie, UPR 9024. 31, Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France.
[3]Department of Biology, University of California San Diego, La Jolla, California, USA.

post-translational modifications[10,11] suggesting that its synthesis follows a strictly regulated pathway. Although several preliminary works hinted at the possible function of MGP as a calcification inhibitor,[3,12] its function has remained unclear until recent gene deletion experiments[13] showed that homozygous mice lacking the MGP gene begin to die 2–3 weeks after birth from calcification of major arteries, which lead to their rupture and exsanguination. Furthermore, mutations in the MGP gene recently have been linked to the human Keutel syndrome,[14] an autosomal recessive disorder characterized by the presence of abnormal cartilage calcifications, midfacial hypoplasia, and pulmonary stenosis. Taken together, these results have shown clearly that "in vivo," MGP plays an important role in the regulation of extracellular matrix calcification and is required for maintenance of cartilage and soft tissue integrity during normal growth and development.

The publication of the sequence for the shark MGP protein[15] has shown that there is a large degree of conservation in those residues considered to be important for the protein function. MGP's posttranslational modifications, such as N- and C-terminal cleavage sites and sites of serine phosphorylation and γ-carboxylation,[11,15] also have been maintained. In addition, MGP also shares significant sequence homology with osteocalcin (or bone Gla protein [BGP]), another vitamin K–dependent Gla-containing protein from the extracellular matrix of bone.[16] Taken together, these results suggest that the function of the protein has remained unaltered despite more than 400 million years of evolution. Given the complexity of the mammalian system, the use of a lower vertebrate as a model system can be advantageous in attempts to relate specific pathways of growth and development with the onset of MGP expression. Features such as high number of progeny and large egg size, external development, and ease in handling both eggs and embryos, together with the fact that it is already an established model for early vertebrate development, makes *Xenopus* a suitable organism to further study the function of MGP.

In this report, we describe the purification of the MGP protein from *Xenopus laevis*, the cloning of its complementary DNA (cDNA) and pattern of tissue distribution, and the expression of MGP during development before metamorphosis and calcification of the skeleton. Considerations on the evolutionary pathway of MGP from fish to man and its relationship with BGP are discussed also.

MATERIALS AND METHODS

Purification of MGP from X. laevis *bone*

Adult *Xenopus* vertebra and long bones were freed from adhering soft tissues and bone marrow, extensively washed in water and acetone, dried, and then powdered using a blender. Bone powder was washed in 6 M guanidine to remove the organic matrix, washed again in water and acetone, air-dried, and kept frozen until used. Extraction was performed from the mineralized bone matrix using a modification of previously described procedures.[15] Briefly, bones were demineralized at 4°C for 4 h with 10% formic acid at a volume to weight ratio of 1 ml acid/10 g of bone. The extracted proteins were dialyzed against water using a 3500-molecular weight cut-off tubing (Spectra/Por; Spectrum Labs, Gardena CA, USA) to precipitate MGP. The precipitated fraction of proteins was collected by centrifugation, resuspended in 6 M guanidine hydrochloride, 0.1 M Tris, pH 9.0, and fractionated by molecular weight over a Sephacryl S-100 HR column (0.9 cm × 150 cm) equilibrated in the same buffer. Fractions of approximately 1.7 ml each were collected and protein content was estimated by spectrophotometer readings at 280 nm and 220 nm. Because incomplete separation was achieved, fractions (24–35) expected to contain MGP were pooled and dialyzed against 50 mM HCl as described previously. The entire content of the dialysis bag was then collected, lyophilized, resuspended in 6 M guanidine hydrochloride, 0.1 M Tris, pH 9.0, and refractionated by molecular weight over a second Sephacryl S-100 HR column (2.0 cm × 150 cm). Fractions of approximately 4.0 ml were collected and protein content was determined as described previously. Fractions 58 and 59 (Fig. 1), thought to contain MGP, were pooled and further analyzed by sodium dodecyl sulfate–poly acrylamide gel electrophoresis (SDS-PAGE).

N-terminal protein sequence analysis

The Sephacryl S-100 HR fractions containing purified *Xenopus* MGP (xMGP) were dialyzed against 50 mM HCl and protein content was estimated by absorbance at 280 nm. Two small aliquots corresponding to 5 μg and 15 μg of total protein (on the assumption that one A280 U equals 1 mg of protein), were freeze-dried, dissolved, and electrophoresed on an 18% SDS/PAGE gel. A major band was present in each case at the expected molecular weight for MGP and no other bands were detected. Therefore, an aliquot of the dialyzed material was adsorbed onto a polyvinylidene difluoride (PVDF) membrane using a ProSpin device (Applied Biosystems Division of Perkin Elmer, Foster City, CA, USA) and then subjected to N-terminal protein sequencing using an Applied Biosystems Model 494 sequenator equipped with an on-line high-performance liquid chromatography (HPLC) for phenylthiohydantoin (PTH)-amino acid detection.

Detection of the phosphorylated serine residues

An aliquot of the MGP peak fraction from the S100 gel filtration column, corresponding to 2 nmol of protein, was adsorbed onto a PVDF membrane using a ProSpin device and rinsed with 20% methanol to remove salts or contaminants. Phosphoserine residues were identified by protein sequence analysis after being converted to *S*-ethylcysteine by reaction with ethanethiol, as described elsewhere.[11,17] The PVDF membrane was placed in a 1.5-ml screw-cap polypropylene tube along with 100 μl of derivatization reagent (stock solution: 80 μl ethanol, 65 μl 5 M NaOH, 60 μl ethanethiol, and 400 μl water) and then flushed with nitrogen and capped. The reaction was allowed to proceed for 2 h at 60°C. Before the start of the sequencer program,

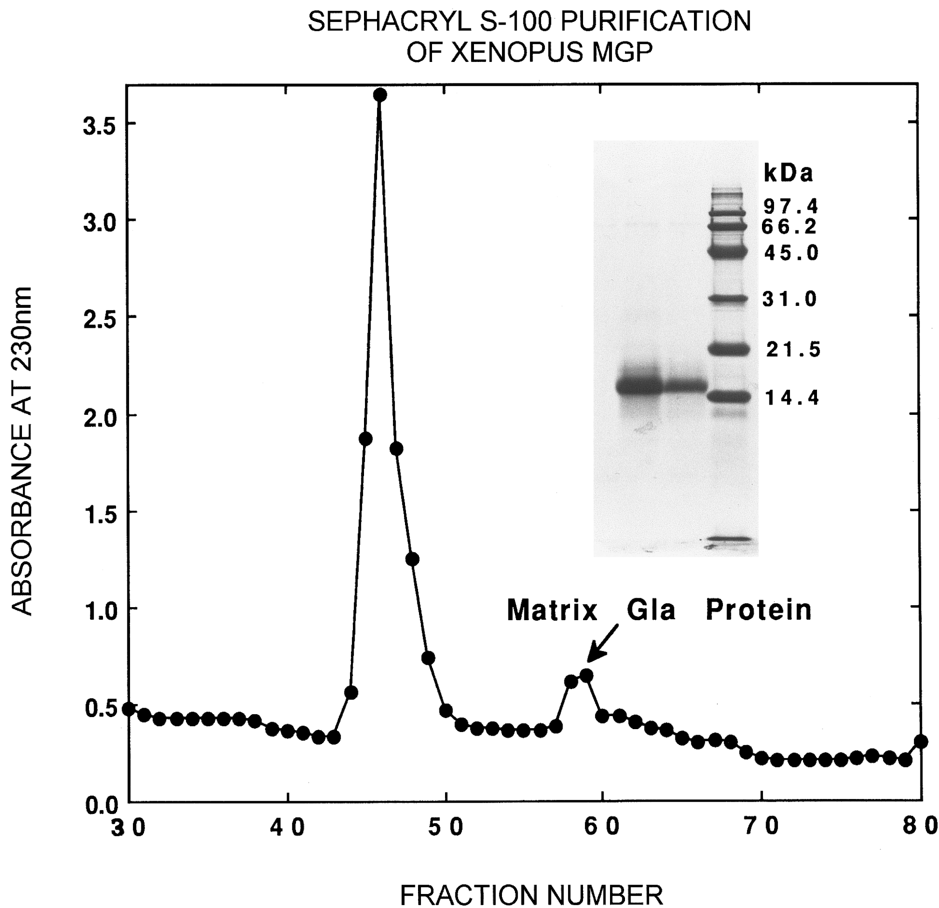

FIG. 1. Purification of MGP from *Xenopus* bone. MGP was extracted from *Xenopus* bone as described in the Materials and Methods section and purified by two sequential Sephacryl S-100 columns. The second Sephacryl S-100 chromatogram is shown. The putative MGP peak (fractions 58–59) was further analyzed by SDS-PAGE. An aliquot of the protein was dialyzed into 50 mM HCl, dried, and two different amounts of the purified protein (15 μg and 5 μg) were loaded onto an 18% SDS/PAGE gel. The gel was stained in 0.1% Coomassie blue and destained as described.[15] BioRad low molecular weight protein standards are in the lane furthest to the right of the gel.

the PVDF membrane was dried for 30 minutes with argon to remove the ethanethiol reagent.[11]

Nucleotide sequence of xMGP cDNA

The N-terminal amino acid sequence of the xMGP protein (26 residues) obtained by protein sequencing was used to construct a forward oligonucleotide primer spanning from residue 1–8 of the mature protein. Total RNA was isolated from *Xenopus* long bones following an established method,[18] reverse-transcribed (RT) using an oligo-d(T) adapter (5′-acgcgtcgacctcgagatcgatg(t)$_{18}$-3′), and amplified by the polymerase chain reaction (PCR) using a degenerate primer designed based on the protein sequence obtained for the xMGP N-terminal region (5′-ta(tc)ga(tc)tcita(ct)ga-(ag)agcca(tc)ga-3′) and a primer corresponding to the adapter sequence (5′-acgcgtcgacctcgagatcgatg-3′). The resulting PCR products were visualized on a 1% agarose gel and the fragment corresponding to the putative xMGP cDNA was cut from the gel, eluted, and cloned into pCR II (Invitrogen, La Jolla, CA, USA). Final identification was achieved by sequence analysis of several specific clones by the dideoxy chain termination method of Sanger.[19] Cloning of the 5′ end of the xMGP cDNA was performed by 5′-rapid amplification of cDNA ends (RACE)–PCR on single-strand cDNA obtained as described previously except that Smart II oligonucleotide (Clontech, Palo Alto, CA, USA) was added in the RT mix. PCR was then performed using a "PCR primer" (SMART PCR cDNA Synthesis Kit; Clontech) specific for the 5′-anchor (from SMART kit) and a MGP-specific reverse primer (5′-ctcttgttgcgctcacgg-3′). The resulting amplified DNA fragments were cloned into the vector pGem-T-easy (Promega, Madison, WI, USA) and further identified by DNA sequence analysis.

Northern blot analysis

Total RNA from several tissues (bone, heart, liver, kidney, gonad, muscle, skin, brain, and intestine) was prepared following an established method.[18] The RNA obtained was then fractionated on a 1.4% formaldehyde-containing agarose gel and transferred onto N^+ nylon membranes (0.45 μm; Nytran; Schleicher & Schuell, Dassel, Germany) by a capillary method.[20] Prehybridization was carried out at 42°C in hybridization solution (Ultrahyb; Ambion, Inc., Austin, TX, USA) for 2–3 h. A specific cDNA probe (spanning from nucleotides 1–570 of the xMGP cDNA) was labeled with [α-^{32}P] deoxycytosine triphosphate (dCTP) using the ReadyprimeII Random Prime Labeling System (Amersham-Pharmacia, Uppsala, Sweden) and separated from unincorporated nucleotides on a MicroSpin S-200HR column (Amersham-Pharmacia). Hybridization was performed overnight under the same conditions described for prehy-

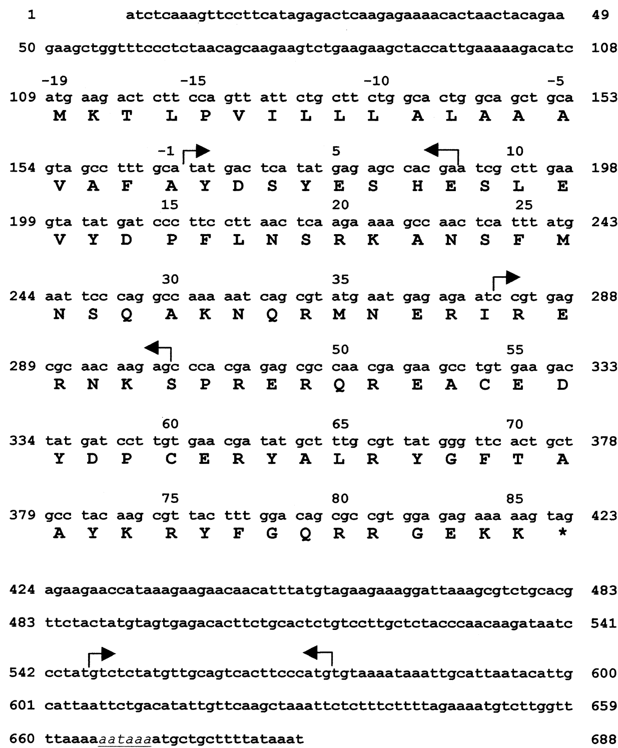

FIG. 2. cDNA and deduced amino acid sequence of the xMGP protein. The cDNA was obtained by RT/PCR and 5′-RACE–PCR amplification of *Xenopus* heart RNA. Numbering on the side is according to the first nucleotide identified as number one in the longest 5′-RACE extension obtained. The arrows indicate the localization of the oligonucleotides used to amplify the MGP cDNA by RT/PCR and 5′-RACE. The first amino acid of the mature protein is identified as +1. The stop codon is indicated by an asterisk and the polyadenylation signal is shown in italics and underlined.

bridization after adding the probe. Blots were washed twice in 2× SSC and 0.1% SDS at 42°C for 5 minutes each and twice in 0.1× SSC and 0.1% SDS at 42°C for 15 minutes each. Autoradiography was performed at −70°C with Kodak XAR-5 film (Amersham-Pharmacia) and Cronex Lightning Plus screens (DuPont, Wilmington, DE, USA).

RT-PCR amplification and Southern hybridization of xMGP message in several developmental stages and tissues of X. laevis

Xenopus specimens were grown at 17°C in our facilities and collected at different times after induced fertilization. One microgram of total RNA extracted from adult tissues (bone, heart, liver, kidney, and muscle) from whole *Xenopus* specimens at various developmental stages (stages 44–58 according to Nieuwkoop and Faber[21]) and from posterior limbs of the froglet stage (stage 65) were RT using an oligo-d(T) adapter (5′-acgcgtcgacctcgagatcgatg(t)$_{18}$-3′) and then one-twentieth of each reaction was amplified by PCR, using two specific oligonucleotide primers designed on the xMGP cDNA sequence previously obtained (forward primer, from nucleotide 248–275; reverse primer, from nucleotide 547–572; Fig. 2 shows localization). PCR reactions were carried for 22 cycles and 30 cycles (one cycle, 30 s at 94°C, 30 s at 60°C, and 30 s at 72°C) followed by a final 10-minute elongation period at 72°C with *Taq* DNA polymerase (Gibco-BRL, Barcelona, Spain). A negative control was made by amplifying with all the reagents including primers but without DNA added to the tube. The resulting PCR products were separated by 1% agarose gel electrophoresis, Southern transferred onto a nylon membrane (Hybond; Amersham), and prehybridized at 42°C for 3 h using 5× SSPE, 10× Denhardt's, 0.5% SDS, and 50 μg/ml heat denatured salmon sperm DNA. Hybridization was performed for 15 h at 42°C in 50% formamide, 6× SSPE, 0.5% SDS, and 50 μg/ml heat denatured salmon sperm DNA and the xMGP cDNA fragment labeled with α-^{32}P dCTP, as described earlier. The membrane was washed for 2 × 15 minutes at room temperature and 2 × 30 minutes at 55°C with 0.1× SSC and 1% SDS and then subjected to autoradiography. As an internal control for the relative amount of RNA used for each sample, the *Xenopus* ornithine decarboxylase (ODC) messenger RNA (mRNA) also was amplified from the same RT reaction. Amplification was performed using two specific primers designed according to the ODC published DNA sequence[22] and was carried out for 30 cycles (one cycle, 30 s at 94°C, 30 s at 60°C, and 30 s at 72°C) followed by a final 10-minute elongation period at 72°C and thus resulting in the amplification of the expected 382-base pair (bp) DNA fragment.

Detection of mineralized structures during Xenopus *development by whole mount staining with Alizarin red/Alcian blue*

Xenopus specimens collected at the same developmental stages used for RT-PCR were fixed overnight at 4°C in 1% freshly made paraformaldehyde and then washed 3 × 10 minutes in Tris buffered saline with triton (TBST) buffer (50 mM Tris, pH 7.4, 150 mM NaCl, and 0.1% Triton X-100) and stored in methanol at 4°C. Whole specimens were hydrated in a decreasing alcohol series and stained for cartilage with Alcian blue 8GX (Sigma, Siutra, Portugal) for various periods of time according to size.[23] Specimens were removed from the staining solution immediately after visible stained structures were observed to avoid decalcification of small structures by the glacial acetic acid–ethanol solution. The specimens were then again hydrated in a six-step decreasing alcohol series and incubated in 1% KOH at room temperature for various periods of time until cleared. Staining with Alizarin red S (Sigma) was performed as with Alcian blue, followed by incubation in various glycerol/1%KOH baths (starting at 1:3 and increasing to 3:1, respectively). Storage in absolute glycerol was initiated only when the specimens were completely clear, with all the internal structures clearly visible. Alkaline-equilibrated phenol (5–10 μl, pH 8) was added to glycerol for prevention of contamination.

Data analysis

The amino acid sequences of the protein were aligned and used to compute a matrix of percentage of differences using

the PHYLIP software package found at http://evolution.genetics.washington.edu/phylip.html.[24] Maximum parsimony and neighbor joining trees (for MGP and BGP sequences) also were constructed with the PHYLIP package. The robustness of the trees generated was tested using 1000 bootstraps.[25] The parsimony approach, which is based on character changes, might help to better understand the evolution of this protein among different organisms and especially the possible loss and gain of some functionally important sites.

RESULTS

Purification and protein sequence analysis of xMGP

MGP was purified from the insoluble fraction of acid-demineralized *Xenopus* bone by gel filtration twice over Sephacryl S-100 HR columns equilibrated in 6 M guanidine HCl and 0.1 M Tris, pH 9.0, as described in the Materials and Methods section. The chromatogram shown in Fig. 1 revealed a small peak eluting at fractions 58–59, corresponding to the elution volume for MGP in this type of column (approximately 240 ml). Only one protein with an apparent molecular weight of 16 kDa was identified in this fraction by SDS/PAGE (Fig. 1, inset) and N-terminal protein sequence analysis confirmed its identity to be MGP. The first 26 amino acids of the mature xMGP obtained by N-terminal sequence analysis were YDXYEXHEXLEVYDPFLNSRKANSFM. This sequence, when compared with the known MGP sequences revealed a higher degree of identity in this region with the mammalian MGPs than with the shark MGP protein previously reported.[15]

Identification of phosphoserine residues in xMGP

The initial sequence analysis of the protein resulted in blanks at positions 3, 6, and 9. Given that all mammalian MGPs analyzed to date have conserved phosphoserine residues at those sequence positions[11] and phosphoserine residues will give blanks unless the protein is modified before sequence analysis,[26] these results suggested that the xMGP protein was phosphorylated in the same residues as its mammalian counterparts. To verify this hypothesis, the xMGP was treated with ethanethiol to convert the putative phosphoserine residues to *S*-ethylcysteines[17] and then subjected to N-terminal sequence analysis. The HPLC separation of the PTH derivative for residues 3, 6, and 9 revealed, in each case, a single peak that eluted just ahead of diphenylthiourea (DPTU; a byproduct of sequencer chemistry) in the chromatograph and coeluted with PTH–*S*-ethylcysteine. Only trace amounts of PTH serine were present at these residues. This indicates that these three residues in xMGP are phosphoserines and that phosphorylation at these positions is nearly complete.

Molecular cloning of xMGP cDNA

The xMGP cDNA was obtained by RT-PCR from total RNA extracted from bone using a primer designed according to the first eight residues of the mature protein sequence.

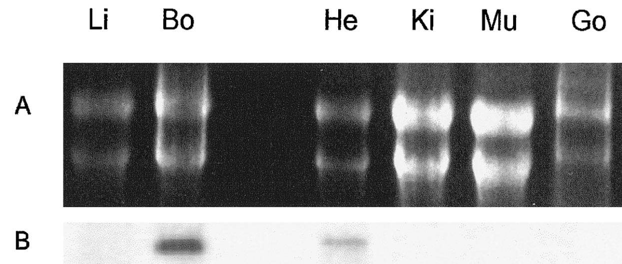

FIG. 3. Analysis of MGP expression in different adult *Xenopus* tissues by Northern blot. Tissue distribution of MGP mRNA was examined by Northern blotting of total RNA extracted from X*enopus* liver (Li), bone (Bo), heart (He), kidney (Ki), muscle (Mu), and gonad (Go). Samples (10–20 μg) were electrophoresed on a 1.4% agarose-formaldehyde gel, transferred to a nylon membrane, and hybridized to a xMGP cDNA as described in the Materials and Methods section. (A) The relative amount of RNA loaded was assessed by the intensity of ethidium bromide staining of ribosomal RNA bands. (B) Positive hybridization using a specific xMGP cDNA probe.

DNA sequence analysis confirmed the identity of the cloned PCR product to be a partial xMGP cDNA, spanning from the first residue of the mature protein to the site of insertion of the poly A tail (Fig. 2). The predicted amino acid sequence was in full agreement with the N-terminal protein sequence obtained for the mature MGP purified from *Xenopus* bone extracts. The 5′ end of the xMGP cDNA was then obtained by 5′-RACE–PCR using a reverse primer defined within the coding sequence, as described in the Materials and Methods section. A DNA fragment was amplified, cloned, and sequenced and found to overlap with the first xMGP cDNA obtained. The full-length sequence of the xMGP cDNA comprises a 5′-untranslated region (5′-UTR) of 108 bp followed by an open reading frame coding for 104 amino acids from the first ATG to the stop codon and 265 bp of 3′-UTR (Fig. 2). By comparison with known MGP cDNAs and in agreement with the N-terminal sequence of the mature xMGP protein, the first 19 residues correspond to the signal peptide followed by an 85-residue mature protein.

Tissue distribution of MGP mRNA

The expression of MGP in adult *Xenopus* tissues was analyzed by Northern hybridization in tissues obtained at dissection of the animal as described in the Materials and Methods section. From all the tissues analyzed, only heart and bone presented a clearly detectable signal for MGP mRNA, ranging from approximately 700 to 800 nucleotides (as deduced from the size of the ribosomal RNAs). This length was compatible with the size of the cloned xMGP cDNA (688 bp) plus the poly A tail (Figs. 2 and 3). No clearly detectable levels for MGP mRNA were seen in extracts of liver, muscle, kidney, or gonad (Fig. 3) or in extracts of skin, brain, and intestine (results not shown). However, when analyzed by RT-PCR using xMGP specific primers, a DNA fragment corresponding to the expected size for the xMGP cDNA also was amplified from kidney, with the highest levels seen in heart and bone. Its identity was confirmed by Southern hybridization with a specific

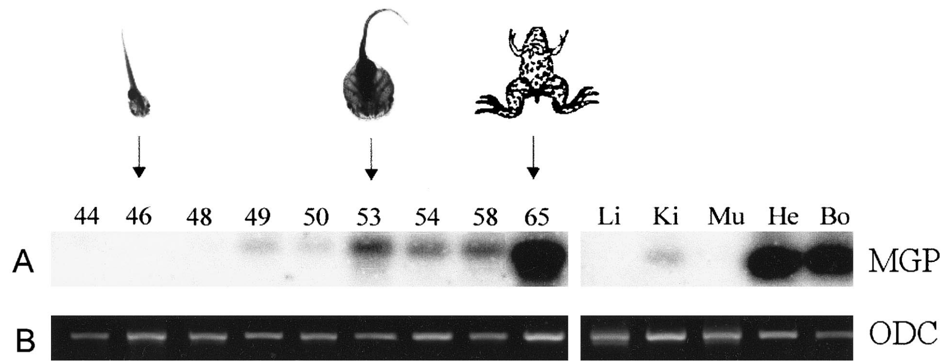

FIG. 4. Analysis of MGP expression in different developmental stages and adult tissues of *Xenopus* by RT/PCR. One microgram of total RNA extracted from *Xenopus* specimens from developmental stage (stages 44–58), from total posterior limbs of froglet (stage 65), and from adult tissues (liver [Li], kidney [Ki], muscle [Mu], heart [He], and bone [Bo]) were subjected to RT-PCR (22 cycles of amplification) using specific xMGP primers. The resulting PCR products were separated by electrophoresis on a 1% agarose gel. Staining with ethidium bromide did not detect the presence of DNA in any lane. The gel was then transferred to a nylon membrane and probed with a specific MGP cDNA as described in the Material and Methods section. (A) Resulting autoradiography. The band corresponds to the expected cDNA fragment amplified from the xMGP mRNA with the forward and reverse primers used. Two independent sets of amplifications from two independent RT reactions were performed and gave comparable results. (B) RT-PCR with primers specific for *Xenopus* ODC gene was performed on the same samples to control for the relative amount of RT mRNA available for PCR and the amplified PCR products were separated by electrophoresis on a 1% agarose gel and stained with ethidium bromide. The external morphology of *Xenopus* specimens at stages 46, 53, and 65 are shown on top of the figure.

xMGP cDNA (Fig. 4) and by DNA sequence analysis after cloning of the PCR products obtained (results not shown).

Expression of MGP mRNA during Xenopus *development*

The presence of xMGP mRNA in *Xenopus* specimens collected through development was analyzed by RT-PCR starting with premetamorphic larval specimens and into metamorphic and postmetamorphic stages. After RT of the RNA samples obtained and using 30 cycles of DNA amplification with MGP-specific primers (see Materials and Methods section for details), a clearly distinguishable PCR product was detected by electrophoresis, beginning with samples collected at stage 46, and shown to hybridize with a labeled xMGP cDNA. A longer exposure of the blot revealed a positive signal even at stage 44 (results not shown), indicating that MGP mRNA was present in all stages of development tested. When the number of amplification cycles was decreased to 22, no distinguishable PCR product was seen in any sample after electrophoresis and ethidium bromide staining. However, hybridization with a specific xMGP probe confirmed the presence of xMGP mRNA in samples from stage 49 onward (Fig. 4A). A longer exposure of the blot revealed a faint signal even at the earlier stages analyzed (results not shown). To prevent the possibility of contamination from genomic DNA, the RT sample was treated with DNaseI (to destroy any remaining genomic DNA in our sample) before PCR amplification. To further exclude the possibility that our positive signal may result from genomic DNA amplification, the primers used were located in both sides of several introns of the xMGP gene (our unpublished data). The use of genomic DNA as template would have resulted in the amplification of a much larger fragment (nearly 4 kilobases [kb]). Our results clearly indicate that MGP mRNA is present during *Xenopus* development, before metamorphosis takes place. Alcian blue/Alizarin red treatment of *Xenopus* specimens collected at the same developmental stages used for RT/PCR confirmed that MGP was detected long before the appearance of the first clearly mineralized structures (stage 54, results not shown).

Primary structure of MGP and evolutionary considerations

The amino acid sequence of the mature xMGP is presented in Fig. 5 along with the amino acid sequences of all other known MGPs and the amino acid sequences of the evolutionarily related BGPs from 11 different species. The sequences are aligned to give maximal homology. The region of maximal homology between MGP and BGP is denoted by a box. There are 24 invariant amino acids among the 8 known MGP sequences (marked in Fig. 5 by a vertical arrow) and 11 invariant amino acids among all the known BGP sequences, including those sequences known but not shown (denoted in Fig. 5 by an arrow). Of these invariant amino acids, eight are common to both MGP and BGP and include two Gla residues and the two Cys residues, which form the disulfide bond.

Analysis by pairwise sequence comparison of mature MGPs from shark to human shows that xMGP shares a higher identity score with mammalian MGPs (55% and 57% identity with human and bovine sequences, respectively) than with the shark MGP (53% identity; Table 1). Nevertheless, from all MGP sequences available to date, xMGP is still the closest neighbor of shark MGP because bovine and chicken MGPs have only 39% and 44%, respectively, identity with the shark sequence (Table 1). When the regions of highest sequence homology between MGP and BGP (Fig. 5,

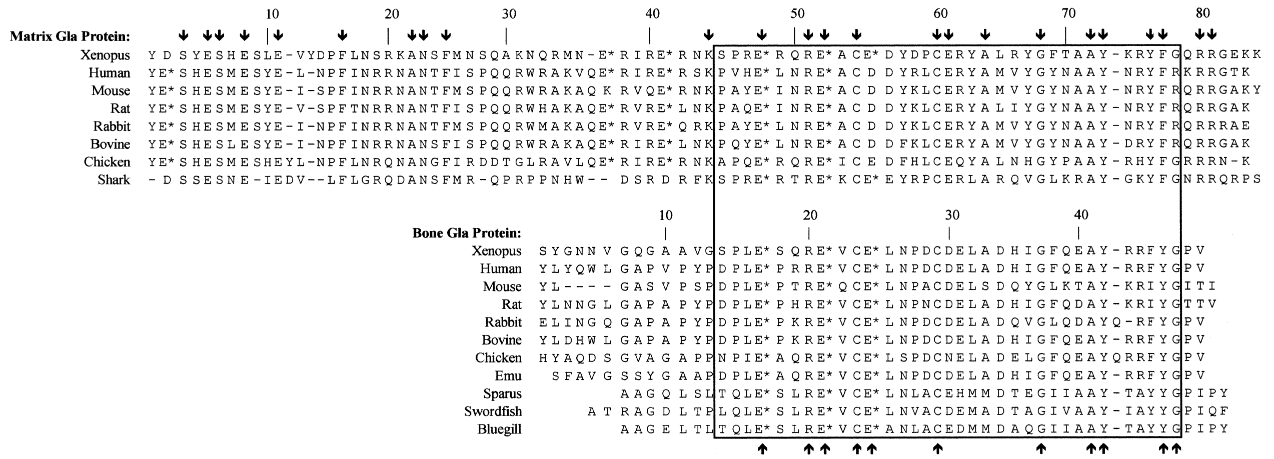

FIG. 5. Protein sequence comparison between mature sequences from all known MGPs and selected BGPs representing different phylogenetic groups. The regions of maximal homology between MGP and BGP are boxed. The positions of γ-carboxyglutamate residues are indicated by (E*). The invariant residues in MGP sequences are denoted by vertical arrows pointing down; the invariant residues in BGP sequences are denoted by vertical arrows pointing up. Dashes indicated gaps in the sequence, introduced to increase homology. Residues are numbered according to the xMGP and BGP proteins. MGP sequences: *Xenopus*, this study; human[38]; mouse[39]; rat[40]; rabbit[41]; cow[42]; chicken[36]; and shark[15] (note, the shark MGP sequence extends beyond what is shown, but this region is omitted because there is no counterpart in any other species). BGP sequences: *Xenopus*[16]; human and mouse[43]; rat[44]; rabbit[45]; cow[46]; chicken[47]; emu[48]; sparus[16]; swordfish[49]; and bluegill.[50]

TABLE 1. PERCENTAGE OF IDENTITY SCORES AMONG MGPS FROM DIFFERENT SPECIES DEDUCED FROM PAIRWISE SEQUENCE ANALYSIS

MGP	*Bovine*	*Chick*	Xenopus	*Shark*	*Genbank accession number*
Human	85	60	55	41	J05572
Bovine	—	59	57	39	GEBOM
Chicken	—	—	56	44	Y13903
Xenopus	—	—	—	53	AF055588.1
Shark	—	—	—	—	AAB31208

Full sequences were analyzed by pairwise sequence comparison as defined in the Materials and Methods section. Results are shown as percentage of identity between the sequences analyzed. Genbank accession numbers for each sequence are shown in the right column.

box) were compared by pairwise identity scores, mammalian MGPs were found to be closer among all BGPs to fish sequences. In addition, shark MGP was closer to nonfish than to fish BGPs (Table 2). Sequence comparisons were completed by drawing phylogenetic trees following neighbor joining (Fig. 6) and maximum parsimony (Fig. 7) approaches on alignment of BGPs and MGPs. Significance of the resulting networks was tested by bootstrapping, as explained in the Materials and Methods section. No clear phylogeny could be inferred inside mammalian MGPs and BGPs but *Xenopus* and chicken MGPs were clearly an intermediate between shark and mammalian MGPs (Fig. 6). Results obtained by maximal parsimony were very similar except that chicken and *Xenopus* sequences appear closer to shark MGP than to the mammalian proteins (one most parsimonious tree, Fig. 7).

DISCUSSION

The purification strategy used to isolate MGP from the long bones of *Xenopus* took advantage of the knowledge that MGPs from other species are insoluble in aqueous, neutral pH buffers and the assumption that xMGP would share this characteristic. This assumption proved to be cor-

TABLE 2. PERCENTAGE OF IDENTITY SCORES IN THE REGION WITH HIGHEST HOMOLOGY BETWEEN MGP AND BGP AMINO ACID SEQUENCES

		BGP sequences		
		Mammals	Xenopus + *Birds*	*Fishes*
MGP sequences	Mammals	29–30	29	32–35
	Xenopus + Chicken	35–44	41–47	38
	Shark	38–49	39–41	35

The amino acids within the boxed region in Fig. 5 were analyzed by pairwise sequence comparison. Sequences were compared as defined in the Materials and Methods section. Results are shown as percentage of identity between the sequences analyzed.

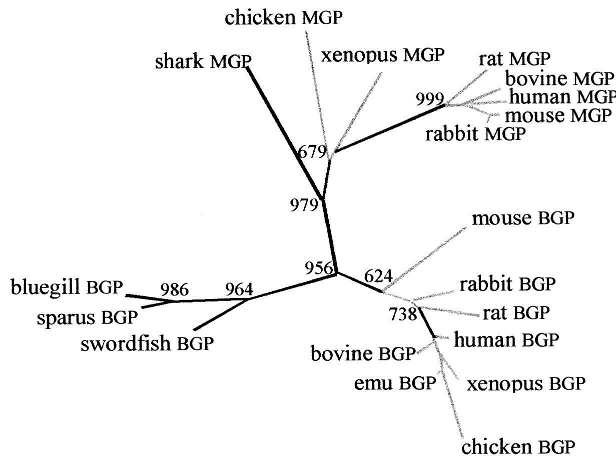

FIG. 6. Comparison of MGP and BGP protein sequences using neighbor joining analysis. Comparison was performed using the sequences shown in Fig. 5. Branch lengths represent evolutionary distances among branching points and are shown in bold when statistically significant. Numbers indicating bootstrap values (1000 resampling) are shown at node except when considered not significant (≤600).

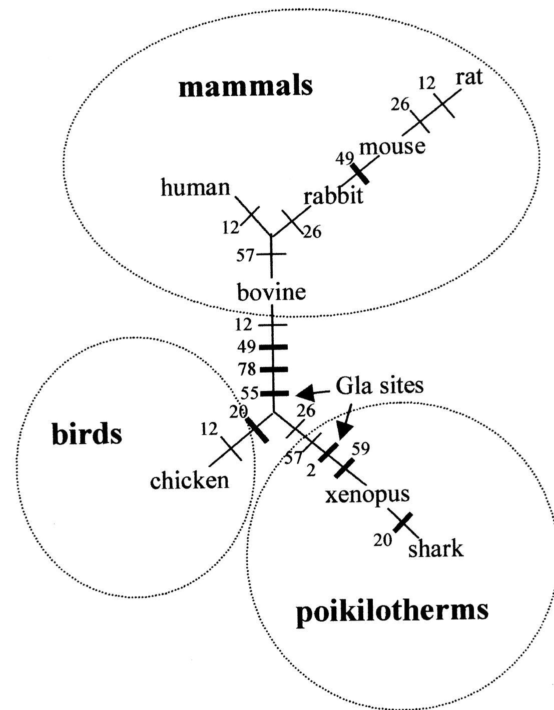

FIG. 7. Analysis of MGP by maximum parsimony analysis. Alignment of MGPs represented in Fig. 5 were analyzed by maximum parsimony. Only one most parsimonious tree was manually obtained. Bar perpendicular to branches represents an amino acid substitution at indicated site according to numbering of Fig. 5 alignment. Note that only sites where each character is shared between at least two of the MGP sequences were considered. Substitutions represented only one time in the tree were indicated by thick bars.

rect. The amount of purified MGP, recovered from the peak fractions shown in the Sephacryl S-100 chromatogram (Fig. 1), was approximately 30 μg/g of *Xenopus* bone. This amount is similar to that previously obtained for bovine bone (60 μg of MGP per gram of bone[15]) and suggests that the deposition of MGP in *Xenopus* bone is comparable with that of mammalian bone.

The N-terminal sequence obtained from the purified mature protein confirmed the presence of high homology between all MGPs in this region and permitted an immediate identification of the purified protein as xMGP. This prompt identification allowed us to proceed rapidly to amplify its cDNA by a combination of RT/PCR and 5′-RACE–PCR and then deduce the complete sequence of the protein.

The full-length xMGP cDNA encodes a polypeptide of 104 amino acids with a signal peptide of 19 residues and a mature protein of 85 residues. The N-terminal first 10 residues contains a repeated motif Ser-Xaa-Glu (SXE) that has been seen in all the MGPs sequenced to date and comprises a serine phosphorylation domain of the protein. The level of phosphorylation determined at serine residues 3, 6, and 9 of xMGP was found to be nearly 100%, which is comparable with shark MGP and, in both cases, significantly higher than in all the mammalian proteins.[11] The justification for this higher level of phosphorylation in MGPs from two lower vertebrates as compared with mammalian MGPs is unclear at the moment but it is likely to be linked to the function and/or regulation of the protein in these organisms. In addition, all MGPs known to date, including the *Xenopus* sequence reported in this article, share a dibasic site at their C terminus (residues 80 and 81 in the *Xenopus* and human proteins; Fig. 5) that has been shown to be a site of proteolytic processing in bovine MGP.[10] The fact that this site is so highly conserved suggests that it is required for the correct regulation of MGP function.

In *Xenopus*, the highest levels of MGP mRNA were found in bones (no efforts were made to remove cartilage from these samples) and heart, in agreement with the tissue distribution of MGP mRNA in mammals. These two tissues have proven to be significantly affected when MGP is either removed or not functional as shown by MGP knockout experiments[13] and in vivo warfarin treatments.[12,27] MGP also was detected in kidney but at levels detectable only by RT-PCR (Fig. 4). This result is in agreement with works performed in mammals that have shown that MGP is present in several different tissues analyzed to date, albeit only in specific cell types.[1–3,28]

The expression of MGP was analyzed in *Xenopus* specimens ranging from feeding premetamorphic larval stage 44 (corresponding to beginning of chondrocranium development) through metamorphic larvae and into postmetamorphic froglet (stage 65). MGP mRNA was detectable within the first days of development in *Xenopus*, long before calcification of the first structures was detected by Alizarin red/Alcian blue histological techniques (stage 54). This suggests that MGP may be required to delay and/or prevent mineralization of soft tissue and cartilage during early development. These results are in agreement with previous data obtained in mammals reporting the expression of MGP early in rat and mouse development.[4,13] Recently, MGP

has been found not only to be a developmentally regulated inhibitor of cartilage mineralization, controlling mineral quantity rather than mineral type, but also to play a role in the regulation of ossification and chondrocyte maturation during early limb development in birds.[29] Our results indicate that in *Xenopus*, MGP expression can be detected before mineralization of cartilage structures takes place, which in our specimens was first detected by Alizarin red/Alcian blue staining at stage 54, in agreement with work previously reported.[30] Because early developmental stages in *Xenopus* are easily accessible due to external fertilization and development and because much is already known concerning the events that take place throughout development in this species, *Xenopus* represents a valuable biological model to investigate the early regulation of gene expression during vertebrate development. Therefore, the availability of MGP cDNA should be useful to further understand the regulation of MGP in vertebrates. Currently, we are determining the gene structure of xMGP so that more in-depth regulatory studies can be performed.

Generally, it is agreed that amphibians were the first vertebrates to move successfully onto land. The transition from fish to crawling four-legged tetrapod occurred about 360 million years ago,[31] by the end of the Devonian period, and implicated the appearance of many adaptations to the new environment such as improved air breathing, increased head mobility, and adaptation to a dry environment. Despite the fact that both bony fish and early amphibians possessed a calcified skeleton, many changes were needed to permit life onto land including development of bone as an active tissue responding to mechanical stresses in the absence of the supporting medium of water. In addition, it required adaptations in handling of salt and water balance and changes in calcium metabolism, for which bone type is critical. In tetrapods, calcium regulation is largely dependent on the interaction between parathyroid gland, which produces PTH, vitamin D, and cellular bone, which acts as a calcium reservoir. In contrast with fish, amphibians have functional parathyroid glands[32] and PTH receptors[33] and a vitamin D receptor[34] has recently been cloned from *Xenopus*. These hormone receptors, involved in calcium homeostasis, are highly homologous to those found in mammals, indicating that amphibians evolved mechanisms for controlling calcium homeostasis very similar to those found in mammals, including establishing bone as a readily available calcium reservoir. Furthermore, amphibians and higher vertebrates have cellular bone whereas fish have both cellular and acellular bone, with prevalence of the latter in higher teleosts.[35] Taken together, these data further stress the importance of amphibians as a model organism to further investigate and understand the function of specific proteins related to calcium metabolism, in particular, those such as MGP, which are involved directly in the regulation of tissue calcification.

Although xMGP was found to have a higher degree of sequence homology with MGPs from mammalian origin than with the shark sequence, it also shared some unique features with shark MGP. Residue 2 of the *Xenopus* and shark mature proteins is an aspartic acid and not a Gla residue as is seen in all the other known MGP sequences (Fig. 5). The xMGP protein has four of its five putative Gla residues in the same positions as the mammalian proteins (Gla residues in positions 37, 41, 48, and 52; Fig. 5) but the fifth Gla residue for xMGP is likely to be residue 55 as is the case for the shark protein. When all the MGPs are aligned with a representative set of the known BGP sequences, this fifth putative Gla residue in *Xenopus* is at a conserved position with a known site of γ-carboxylation in BGPs (Fig. 5). This feature is suggestive of the common ancestor from which both MGP and BGP are most likely derived and indicates that mammalian MGPs have diverged from all other known MGP sequences at this position.

When comparing the *Xenopus* with the mammalian MGP sequences, two substitutions in conserved sites occur in xMGP (positions 19 and 31), all within the putative γ-carboxylase recognition site. These divergences add to those seen within the same protein domain for the shark MGP sequence[15] and the chicken MGP[36] and confirm that this is one of the most variable regions of MGP. In the human MGP, this sequence was found to share some homology with a region of the mammalian γ-carboxylase enzyme,[37] and it was hypothesized that changes within this region may be linked to the degree to which MGP is γ-carboxylated by this enzyme. This process must involve recognition mechanisms different from those used for the other known vitamin K–dependent proteins, which, in contrast with MGP, are all synthesized as propeptide-containing precursors. Because complete sequences for the carboxylase enzyme from species other then mammals are not available, at the present time it is not possible to ascertain whether an identical correlation is seen in *Xenopus* or in other nonmammalian species between this domain of MGP and the substrate binding domain of the γ-carboxylase, as reported for the mammalian proteins.[37]

In contrast to the changes seen within the γ-carboxylation recognition domain, the invariant residues located within the phosphorylation domain (the SXE motif within the N-terminal first 10 amino acids) and the Gla-containing domain (Gla-Xaa-Xaa-Xaa-Gla-Xaa-Cys sequence, from amino acid 48–54), previously observed for mammalian and shark MGPs,[15] also are invariant when MGPs from *Xenopus* and chicken are analyzed. This supports the hypothesis that these residues are important for the full activity of this protein. Some of these residues also are invariable when MGP and BGP sequences are compared,[16] suggesting that their location within the protein structure may be important for the correct folding and/or function of MGP and BGP and further supports the hypothesis stated previously.

Although the close proximity found between mammalian MGPs and fish BGPs, as well as between shark MGP and nonfish BGPs, seems intriguing, a possible explanation for this may be simply the result of different evolutionary histories among BGP and MGP and perhaps different evolution rates among fish and nonfish BGPs. To better ascertain this possibility, it would be necessary to have available additional sequences of MGPs from several bony fishes as intermediates to shark and amphibian MGPs.

ACKNOWLEDGMENTS

The authors thank Dr. Glen Sweeney and Dr. Stavroula Kousteni for providing the *Xenopus* ODC primers (GS) and helpful discussions, Dr. Didier Aurelle for critical comments on the evolutionary issues, and Natercia Conceição and Paulo Gavaia for expert technical assistance on obtaining the 5′ end of the xMGP cDNA and Alizarin red/Alcian blue staining of the collected specimens, respectively. This work was partially funded by NATO/CRG940751/SA5.2.05, Praxis XXI/BIA 469/94, and the National Institutes of Health (NIH; grant AR 25921). M.C.P.O. and J.P.R. were the recipients, respectively, of a postdoctoral (Praxis XXI/BPD/18816) and initiation to research (Praxis XXI/BICJ-2985) fellowship from the Portuguese Science and Technology Foundation.

REFERENCES

1. Fraser JD, Price PA 1988 Lung, heart, and kidney express high levels of mRNA for the vitamin K-dependent matrix Gla protein. J Biol Chem **23:**11033–11036.
2. Cancela ML, Price PA 1992 Retinoic acid induces matrix Gla protein gene expression in human cells. Endocrinology **130:** 102–108.
3. Hale JE, Fraser JD, Price PA 1988 The identification of matrix Gla protein in cartilage. J Biol Chem **263:**5820–5824.
4. Otawara Y, Price PA 1986 Developmental appearance of matrix Gla protein during calcification in the rat. J Biol Chem **261:**10828–10832.
5. Cancela ML, Williamson MK, Price PA 1993 Retinoic acid increases matrix Gla protein in rat plasma. Nutr Res **13:**87–91.
6. Kirfel J, Kelter M, Cancela ML, Price PA, Schule R 1997 Identification of a novel negative retinoic acid responsive element in the promoter of the human matrix Gla protein gene. Proc Natl Acad Sci USA **94:**2227–2232.
7. Fraser JD, Otawara Y, Price PA 1988 1,25-Dihydroxyvitamin D_3 stimulates the synthesis of matrix Gla protein by osteoasarcoma cells. J Biol Chem **263:**911–916.
8. Cancela ML, Hu B, Price PA 1997 Effect of cell density and growth factors on matrix Gla protein expression by normal rat kidney cells. J Cell Physiol **171:**125–134.
9. Zhao J, Warburton D 1997 Matrix Gla protein gene expression is induced by transforming growth factor-beta in embryonic lung culture. Am J Physiol **273:**282–287.
10. Hale JE, Williamson MK, Price PA 1991 Carboxy terminal processing of matrix Gla protein. J Biol Chem **266:**21145–21149.
11. Price PA, Rice JS, Williamson MK 1994 Conserved phosphorylation of serines in the Ser-X-Glu/Ser(P) sequences of the vitamin K-dependent matrix Gla protein from shark, lamb, rat, cow and human. Protein Sci **3:**822–830.
12. Price PA, Williamson MK, Haba T, Dell RB, Jee WSS 1982 Excessive mineralization with growth plate closure in rats on chronic warfarin treatment. Proc Natl Acad Sci USA **79:**7734–7738.
13. Luo G, Ducy P, Mckee MD, Pinero GJ, Loyer E, Behringer RR, Karsenty G 1997 Spontaneous calcification of arteries and cartilage in mice lacking matrix Gla protein. Nature **386:**78–81.
14. Munroe PB, Olgunturk RO, Fryns JP, Van Mandelgem L, Ziereisen F, Yuksel B, Gardiner RM, Chung E 1999 Mutations in the gene encoding the human matrix Gla protein cause Keutel syndrome. Nat Genet **21:**142–144.
15. Rice JS, Williamson MK, Price PA 1994 Isolation and sequence of the vitamin K-dependent matrix Gla protein from the calcified cartilage of the Soupfin shark. J Bone Miner Res **9:**567–576.
16. Cancela ML, Williamson MK, Price PA 1995 Amino-acid sequence of bone Gla protein from the African clawed toad *Xenopus laevis* and the fish *Sparus aurata*. Int J Pept Protein Res **46:**419–423.
17. Meyer HE, Hoffmann-Posorske E, Heilmeyer LMG Jr 1991 Determination and location of phosphoserine in proteins and peptides by conversion to *S*-ethylcysteine. Methods Enzymol **201:**169–185.
18. Chomczynski P, Sacci N 1987 Single step method of RNA isolation by the acid-guanidinium-thiocyanate phenol-chloroform extraction. Anal Biochem **162:**156–159.
19. Sanger F, Nicklen S, Coulson AR 1977 DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA **74:** 5463–5467.
20. Sambrook J, Fritsch EF, Maniatis T 1989 Extraction, purification and analysis of messenger RNA from eukaryotic cells. In: Ford N, Nolan C, Ferguson M (eds.) Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, pp. 7.37–7.52.
21. Nieuwkoop PD, Faber J 1967 Normal table of *Xenopus laevis* (Daudin). In: Nieuwkoop PD, Farber J (eds.) A Systematical and Chronological Survey of the Development from the Fertilized Egg Till the End of Metamorphosis, 2nd ed. North-Holland Publishing Company, Amsterdam, The Netherlands, pp. 162–188.
22. Bassez T, Paris J, Omilli F, Dorel C, Osbourne HB 1990 Post-transcriptional regulation of ornithine decarboxylase in *Xenopus laevis* oocytes. Development **110:**955–962.
23. Gavaia PJ, Sarasquete MC, Cancela ML 2000 Detection of mineralized structures in very early stages of development of marine teleostei using a modified Alcian blue-Alizarin red double staining technique for bone and cartilage. Biotech Histochem **75:**79–84.
24. Felsenstein J 1989 Phylogeny inference package (version 3.2). Cladistics **5:**164–166.
25. Felsenstein J 1985 Confidence limits on phylogenies: An approach using the bootstrap. Evolution **39:**783–791.
26. Meyer HE, Hoffman-Posorske E, Korte H, Heilmeyer LMG Jr 1986 Sequence analysis of phosphoserine-containing peptides. FEBS Lett **204:**61–66.
27. Price PA, Faus SA, Williamson MK 1998 Warfarin causes rapid calcification of the elastic lamellae in rat arteries and heart valves. Arterioscler Thromb Vasc Bio **18:**1400–1407.
28. Rannels SR, Cancela ML, Wolpert EB, Price PA 1993 Matrix Gla protein mRNA expression in cultured type II pneumocytes. Am J Physiol **265:**L270–L278.
29. Yagami K, Suh J-Y, Enomoto-Iwamoto M, Koyama E, Abrams WR, Shapiro IM, Pacifici M, Iwamoto M 1999 Matrix Gla protein is a developmental regulator of chondrocyte mineralization and, when constitutively expressed, blocks endochondral and intramembranous ossification in the limb. J Cell Biol **147:**1097–1108.
30. Trueb L, Hansen J 1992 Skeletal development in *Xenopus laevis* (Anura: Pipidae). J Morphol **214:**1–41.
31. Strickberger MW 2000 From air to water: Amphibians, reptiles, and birds. In: Evolution, 3rd ed. Jones and Bartlett Publishers, Sudbury, MA, USA, pp. 409–441.
32. Scrivastan AK, Das VK, Das S, Sasayama Y, Suzuki N 1995 Amphibian parathyroids: Morphological and functional aspects. Microsc Res Tech **32:**79–90.
33. Bergwitz C, Klein P, Kohno H, Forman SA, Lee K, Rubin D, Juppner H 1998 Identification, functional characterization, and developmental expression of two nonallelic parathyroid hormone (PTH)/PTH-related peptide receptor isoforms in *Xenopus laevis* (Daudin). Endocrinology **139:**723–732.
34. Li YC, Bergwitz C, Juppner H, Demay MB 1997 Cloning and characterization of the vitamin D receptor from *Xenopus laevis*. Endocrinology **138:**2347–2353.

35. Parenti LR 1986 The phylogenetic significance of bone types in euteleost fishes. Zool J Linn Soc **87:**37–51.
36. Wiedemann M, Trueb B, Belluoccio D 1998 Molecular cloning of avian matrix Gla protein. Biochem Biophys Acta **1395:** 47–49.
37. Price PA, Williamson MK 1993 Substrate recognition by the vitamin K-dependent gamma-glutamyl carboxylase: Identification of a sequence homology between the carboxylase and the carboxylase recognition site in the substrate. Protein Sci **2:**1987–1988.
38. Cancela L, Hsieh C-L, Francke U, Price PA 1990 Molecular structure, chromosome assignment and promoter organization of the human matrix Gla protein gene. J Biol Chem **265:** 15040–15048.
39. Ikeda T, Yamaguchi A, Icho T, Tsuchida N, Yoshiki S 1991 cDNA and deduced amino acid sequence of mouse matrix Gla protein: One of five glutamic acid residues potentially modified to Gla is not conserved in the mouse sequence. J Bone Miner Res **6:**1013–1017.
40. Price PA, Fraser JD, Metz-Virca G 1987 Molecular cloning of matrix Gla protein: Implications for substrate recognition by the vitamin K-dependent gamma carboxylase. Proc Natl Acad Sci USA **84:**8335–8339.
41. Sohma Y, Suzuki T, Sasano H, Nagura H, Nose M, Yamamoto T 1994 Expression of mRNA for matrix gamma-carboxyglutamic acid protein during progression of atherosclerosis in aorta of Watanabe heritable hyperlipidemic rabbits. J Biochem **116:**747–751.
42. Price PA, Williamson MK 1985 Primary structure of bovine matrix Gla protein, a new vitamin K-dependent bone protein. J Biol Chem **260:**14971–14975.
43. Celeste AJ, Rosen V, Buecker JL, Kriz R, Wang EA, Wozney JM 1986 Isolation of the human gene for bone Gla protein utilizing mouse and rat cDNA clones. EMBO J **5:**1885–1890.
44. Pan LC, Price PA 1985 The propeptide of rat bone gamma-carboxyglutamic acid protein shares homology with other vitamin K-dependent protein precursors. Proc Natl Acad Sci USA **82:**6109–6113.
45. Virdi AS, Willis AC, Hauschka PV, Triffitt JT 1991 Primary amino acid sequence of rabbit osteocalcin. Biochem Soc Trans **19:**374S.
46. Price PA, Poser JW, Raman N 1976 Primary structure of the gamma-carboxyglutamic acid-containing protein from bovine bone. Proc Natl Acad Sci USA **73**:3374–3375.
47. Carr SA, Hauschka PV, Biemann K 1981 Gas chromatographic mass spectrometric sequence determination of osteocalcin, a gamma-carboxyglutamic acid-containing protein from chicken bone. J Biol Chem **256:**9944–9950.
48. Huq NL, Tseng A, Chapman GE 1987 The amino acid sequence of emu osteocalcin. Gas phase sequencing of Gla-containing proteins. Biochem Int **15:**271–277.
49. Price PA, Otsuka AS, Poser JW 1977 Comparison of gamma-carboxyglutamic acid-containing proteins from bovine and swordfish bone: Primary structure and Ca^{2+} binding. In: Wasserman RH, Corradino RA, Carafoli E, Kretsinger RH, MacLennan DH, Siegel FL (eds.) Calcium Binding Proteins and Calcium Function. Elsevier North-Holland, Amsterdam, The Netherlands, pp. 333–337.
50. Nishimoto SK, Araki N, Robinson FD, Waite JH 1992 Discovery of bone gamma-carboxyglutamic acid protein in mineralized scales: The abundance and structure of *Lepomis macrochirus* bone gamma-carboxyglutamic acid protein. J Biol Chem **267:**11600–11605

Address reprint requests to:
Leonor Cancela, Ph.D.
Centro de Ciências do Mar
Universidade do Algarve-UCTRA
Campus de Gambelas, 8000–810 Faro, Portugal

Received in original form July 10, 2000; in revised form March 14, 2001; accepted March 19, 2001.