Biochem. J. (1992) 281, 513–517 (Printed in Great Britain)

Molecular analysis of *Xenopus laevis* SPARC (Secreted Protein, Acidic, Rich in Cysteine)

A highly conserved acidic calcium-binding extracellular-matrix protein

Sashko DAMJANOVSKI, Fina LIU and Maurice RINGUETTE*
Department of Zoology, University of Toronto, Toronto, Ontario, Canada M5S 1A1

SPARC (Secreted Protein, Acidic, Rich in Cysteine) is expressed as a 1.6 kb mRNA in *Xenopus laevis*. On the basis of cDNA sequence analysis, *Xenopus* SPARC has a core M_r of 32643, with one potential *N*-glycosylation site. Western analysis of SPARC isolated from *Xenopus* long bone indicates that the mature protein has an M_r of 43000. At the amino acid level, *Xenopus* SPARC has 78–79% sequence similarity to mouse, bovine and human SPARC. The least-conserved region is found within the *N*-terminal glutamic acid-rich domain, with the *C*-terminal Ca^{2+}-binding domain being the most conserved. Adult *Xenopus* tissues show the same pattern of tissue-specific distribution of SPARC mRNAs as adult mouse.

INTRODUCTION

The extracellular matrix (ECM) represents a complex network of highly specialized macromolecules. Analyses of the structural and binding properties of ECM components such as collagens, proteoglycans and fibronectins have provided valuable insight into some of the underlying mechanisms governing cell migration, morphogenesis and differentiation (McDonald, 1988). Variations in tissue distribution and isoforms of individual ECM components in part contribute to the highly specialized forms and activities of different tissues. The present investigation was focused on characterizing the structural properties of *Xenopus* SPARC (Secreted Protein, Acidic, Rich in Cysteine; osteonectin/BM-40), a Ca^{2+}-binding ECM glycoprotein.

SPARC was first identified as a major non-collagenous component of bovine subperiosteal bone (Termine *et al.*, 1981). However, recent studies have demonstrated that SPARC is also widely distributed in connective tissues and basement membranes of embryonic and adult soft tissues undergoing rapid growth, cellular differentiation and morphological changes (e.g. endodermal, epidermal and extraembryonic tissues) (Wasi *et al.*, 1984; Holland *et al.*, 1987).

The physiochemistry of SPARC indicates that it may regulate calcium phosphate crystal growth in both mineralized and non-mineralized tissues. SPARC is a potent inhibitor of hydroxyapatite-seeded crystal growth *in vitro* (Romberg *et al.*, 1986). The level of inhibition is similar in the presence or absence of collagen, despite SPARC having a selective affinity for different collagen types (Sage *et al.*, 1989).

Other studies indicate that SPARC has broader or more complex function(s), unrelated to regulating tissue mineralization. Tissue cultures of bovine endothelial cells, fibroblasts or smooth-muscle cells round up and stop migrating in the presence of SPARC, suggesting a role for the protein in cell migration and morphogenesis (Sage *et al.*, 1989). A role in regulating haemostasis is also implicated by (1) its affinity for thrombospondin and (2) its presence in, and secretion by, thrombin-activated platelets (Kelm & Mann, 1990).

To shed further light on the role of SPARC, we sought to examine the degree of conservation of SPARC sequences between amphibians and mammals. The high conservation of amino acid sequences and pattern of mRNA expression are consistent with SPARC having evolutionarily conserved function(s) between amphibians and mammals.

MATERIALS AND METHODS

Northern blotting

Total cellular RNA was extracted from liver, lung, muscle, heart and kidney tissues of adult CD-1 mice and adult *Xenopus laevis*. RNA was prepared by the guanidinium thiocyanate/CsCl method, separated in denaturing agarose/formaldehyde gels and transferred to Duralose membrane (Stratagene) (Sambrook *et al.*, 1989). RNA was u.v.-cross-linked to the membrane with a Stratagene Stratalinker. Blots containing 10 μg of total RNA per lane were prehybridized [50% (v/v) formamide/5 × SSC (1 × SSC is 0.075 M-NaCl/0.05 M-sodium citrate)/5 × Denhardt's solution/0.05 M-Tris (pH 7.6)/0.5% SDS/0.04 M-EDTA/tRNA (100 μg/ml)] at 37 °C for 1 h. cDNA representing full-length *Xenopus laevis* SPARC (see Fig. 2 below) or mouse SPARC (a gift from Dr. Brigid Hogan, Vanderbilt University, Nashville, TN, U.S.A.) was ^{32}P-labelled by using a random-primer-labelling kit (BRL). Labelled [(5–10) × 10^8 c.p.m.] and denatured probe was added to fresh hybridization buffer, added to the blot, and incubated overnight at 37 °C. Blots were washed at moderately high stringency (twice with 2 × SSC at room temperature, twice with 1 × SSC at 50 °C, and once with 0.5 × SSC at 50 °C, all with 0.1% SDS and for 20 min each) and exposed to Kodak X-Omat film overnight at −70 °C.

Isolation and sequence analysis of *Xenopus* SPARC cDNAs

Full-length mouse SPARC cDNA was used to screen 500000 clones (Sambrook *et al.*, 1989) of a *Xenopus laevis* stage-24 cDNA library constructed in λgt11 (a gift from Dr. T. Sargent, National Institutes of Health, Bethesda, MD, U.S.A.). Tertiary screenings identified four unique clones. Cloned fragments, identified by hybridization to labelled mouse SPARC probe after Southern transfer, were excised and subcloned into the multi-

Abbreviations used: SPARC, Secreted Protein, Acidic, Rich in Cysteine; ECM, extracellular matrix; 1 × SSC, 0.075 M-NaCl/0.05 M-sodium citrate; PBS, phosphate-buffered saline (for composition, see the text); TBS, Tris-buffered saline (for composition, see the text).

* To whom correspondence should be addressed.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession no. X62483.

cloning site of the twin promoter plasmid pGEM4Z (Pharmacia) and transformed into competent DH5 alpha cells (Sambrook *et al.*, 1989). Exonuclease III deletions were performed (Henikoff, 1987) to achieve deletions that differed by about 200 bp. Double-stranded DNA was prepared for sequencing as described in Promega's (1989) alkali-denaturation protocol. Purified plasmid DNA was then sequenced according to the protocol given for the U.S. Biochemical Sequenase kit (using [^{35}S]ATP). All regions were sequenced at least twice (up to five times) in each direction. Buffer-gradient 6%-(w/v)-polyacrylamide sequencing gels were run and handled as described by Sambrook *et al.* (1989).

Isolation and Western analysis of *Xenopus* SPARC

The isolation of SPARC from *Xenopus laevis* bone was as described by Domenicucci *et al.* (1988), with the following modifications. Femurs and tibia/fibulas were collected from 40 adult *Xenopus* males and females. Crushed bones were placed in 50 ml of phosphate-buffered saline (PBS; 110 mM-NaCl/10 mM-phosphate, pH 7.4) with proteinase inhibitors and extracted seven times for 24 h with continuous shaking for each extraction (Domenicucci *et al.*, 1988). The tissues were then similarly extracted seven times with G buffer (4 M-guanidinium chloride/50 mM-Tris, pH 7.4, containing proteinase inhibitors). The tissues were then washed three times, for 24 h each, with the PBS buffer. Subsequently seven extractions were performed as described above with E buffer (0.5 M-EDTA/50 mM-Tris, pH 7.4, containing proteinase inhibitors). The tissues were rinsed with the PBS buffer between subsequent E buffer extractions (E extract). All the EDTA extractions were pooled, dialysed against 10 mM-ammonium bicarbonate and freeze-dried.

F.p.l.c. of the *Xenopus* bone E extract was performed on a Pharmacia system with a Fast Q ion-exchange column. Fractions were dialysed against 10 mM-ammonium bicarbonate and then freeze-dried.

Western blots were prepared as described by Towbin *et al.* (1979). Anti-(bovine SPARC) serum raised in rabbits (obtained from Dr. Larry Fisher, National Institutes of Health, Bethesda, MD, U.S.A. was diluted 1:30 in Tris-buffered saline (TBS; 50 mM-Tris/150 mM-NaCl, pH 7.4) containing 0.5 mg of BSA/ml and 0.1% NaN_3. The alkaline phosphatase-conjugated secondary antibody was localized with the Nitro Blue Tetrazolium chloride/5-bromo-4-chloroindol-3-yl phosphate reaction.

Secondary-structure analysis

The sequence of the cDNA clone was used to determine the amino acid sequence. The amino acid sequence was then used to analyse the secondary structure. A hydropathicity profile was constructed according to the Kyte–Doolittle (1982) algorithm. α-Helices, β-sheets and β-turns were approximated as described by Chou & Fasman (1978). Other structural features of the sequence were determined by using P.C. Gene Software (Intelligenetics, Geneva, Switzerland).

RESULTS

Pattern of SPARC mRNA expression in *Xenopus*

We recently demonstrated that RNA sequences, which cross-hybridize to a mouse SPARC cDNA probe, are detected in vertebrates spanning an evolutionary diverse group of organisms, ranging from agnatha (jawless fish) to mammals (Ringuette *et al.*, 1991). Hence the mouse cDNA was used as a probe to screen a *Xenopus* stage-24 cDNA library. Several *Xenopus* cDNA clones were found to hybridize to a 1.6 kb polyadenylated RNA transcript isolated from *Xenopus* tissues and to its cognate 2.2 kb SPARC transcript in mouse (Figs. 1*a* and 1*b* respectively).

Even though amphibians and mammals diverged 300 million years ago, Northern analysis indicated the tissue-specific distribution of *Xenopus* SPARC mRNAs (Fig. 1*a*) was similar to that of mouse SPARC mRNAs (Fig. 1*b*) in adults. In both organisms, highest levels of SPARC mRNA were found in lung tissues. Kidney and heart tissues showed similar intermediate levels of SPARC mRNAs, with lowest levels of expression observed in muscle and liver tissues.

Xenopus SPARC sequence and comparison with mammalian SPARC

Sequence analysis revealed that the longest cDNA insert (7B-1, 1046 bp) overlapped a shorter (15-1, 474 bp) clone which together represent nearly the full-length RNA sequence. Fig. 2 shows the coding region of *Xenopus* SPARC cDNA and a comparison of its derived amino acid sequence with that of mouse, bovine and human SPARC. The AUG start codon and the codon coding for the second amino acid were not represented in the *Xenopus* SPARC cDNA clones (as compared with the three other known SPARC sequences) (Fig. 2). For clarity, and to emphasize differences, the only amino acids shown for the mammalian species reflect differences relative to *Xenopus* SPARC. cDNA sequence analysis revealed that there was a 900-nucleotide open reading frame coding for 300 amino acids (Fig. 2). On the basis of three criteria we postulated that the first 17 amino acids represent the leader peptide, leaving a 283-amino-acid (32643 Da) mature core protein product: (1) the *N*-terminal regions of the mammalian SPARC sequences are known, and there is 100% complete sequence conservation of the *N*-terminus of the mature protein in all the mammalian SPARC sequences (Fig. 2); (2) the putative 17-amino-acid leader sequence is highly hydrophobic (see Fig. 3), a hallmark of leader-peptide sequences; and (3) a consensus signal-peptide cleavage site was identified between amino acids 17 and 18 (P.C. Gene Software). A single TAA termination exists, followed by a 646-bp 3′ untranslated region terminating with a AATAAA consensus polyadenylation site.

The amino acid difference matrix shown in Table 1 demonstrates that *Xenopus* SPARC has 78–79% sequence identity with mouse, bovine and human SPARC. When conservative amino acid substitutions and reversal of adjacent amino acids were taken into consideration, the sequence similarity to mammalian SPARC was demonstrated to be 90–91%.

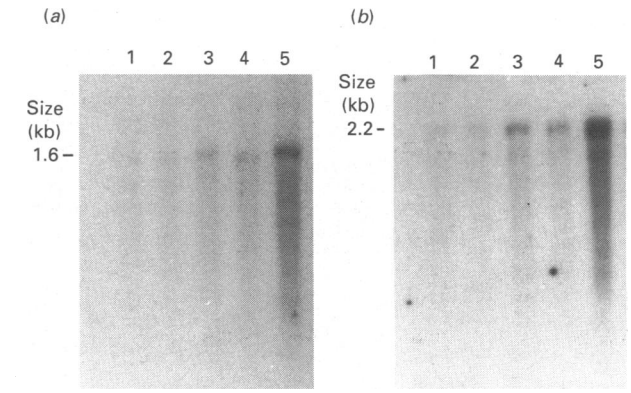

Fig. 1. Expression of SPARC RNA in *Xenopus* and mouse

Total RNA (10 μg) extracted from *Xenopus* and mouse muscle (1), liver (2), kidney (3), heart (4) and lung (5) tissues were fractionated by formaldehyde/agarose-gel electrophoresis, Northern-blotted on to Duralose, and probed with ^{32}P-labelled full-length *Xenopus* (*a*) and mouse (*b*) SPARC cDNA probes. (*a*) Represents *Xenopus* RNA and (*b*) represents mouse RNA.

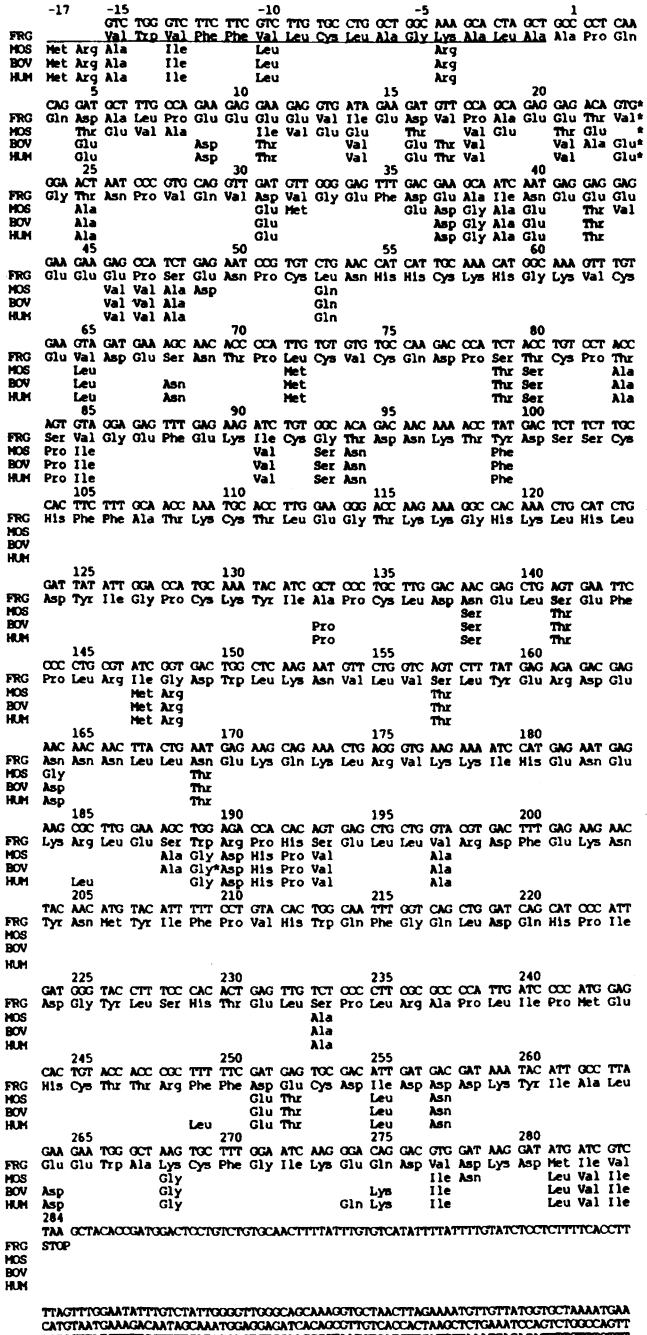

Fig. 2. RNA sequence and predicted amino acid sequence of a *Xenopus laevis* SPARC: comparisons with mouse (MOS), bovine (BOV) and human (HUM) SPARC

A proposed signal peptide is underlined. The secreted core protein of 32634 Da is predicted from the single open reading frame. A single putative *N*-linked glycosylation site (Asn-Xaa-Thr/Ser) is present at amino acid 96. The polyadenylation signal AATAAA is present just upstream of the 3′ end. Of the mouse, bovine and human SPARC sequences, only changes in amino acid sequence relative to *Xenopus* (FRG) SPARC are shown. The asterisk between amino acids 23 and 24 indicates that there is a deletion of three amino acids in the *Xenopus* SPARC protein. The mouse and bovine SPARCS contain a Val-Pro-Val sequence, whereas the sequence Val-Ser-Val is found in human SPARC. The asterisk at amino acid positions 189/190 reflects the presence of an additional cysteine residue in bovine SPARC.

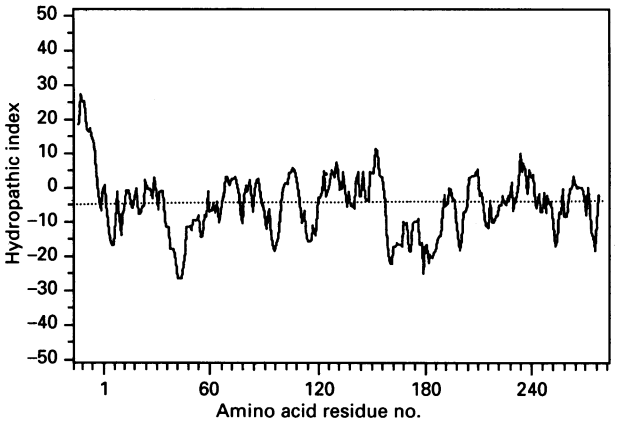

Fig. 3. Hydropathic profile of *Xenopus* SPARC

The hydropathic inclination of SPARC was determined by the Kyte & Doolittle (1982) algorithm. Spans of seven amino acids were used to determine the hydropathic value of a given amino acid side chain.

Table 1. Difference scores for SPARC amino acid sequences

Values in the upper-right triangle are for percentage amino acid conservation; those in the lower-left triangle, shown in **bold**, are for percentage conservation after elimination of conservative amino acid substitutions.

	Frog	Mouse	Bovine	Human
Frog		78	79	79
Mouse	**91**		92	90
Bovine	**90**	**95**		98
Human	**90**	**95**	**99**	

Structural analysis of *Xenopus* SPARC

With the exception of the putative leader-peptide sequence, the hydropathic profile (Fig. 3) revealed that *Xenopus* SPARC was punctuated with several major hydrophilic segments. No archetypical hydrophobic membrane-spanning segments were evident within the mature protein, consistent with SPARC being a globular extracellular protein. However, several small hydrophobic regions exist, in particular within the central region of the sequence, that could play a critical role in the tertiary folding of SPARC into a globular structure. The secondary structure predicted by the Chou & Fasman (1978) analysis indicated that *Xenopus* SPARC is 41% α-helix, 21% β-sheet, 21% β-turn and 17% random coil (Fig. 4).

Sequence comparison between frog and mammalian SPARC (Fig. 2) revealed that the acidic *N*-terminal glutamic acid-rich domain (amino acids 1–50) is the least conserved region of SPARC. The *Xenopus* SPARC sequence demonstrated two repeats of glutamic acid residues: a tetramer at position 9, and a hexamer at position 41. The cysteine-rich domain (amino acids 51–135) showed approx. 50% greater sequence conservation than the glutamic acid-rich *N*-terminal domain. The 11 cysteine residues found within this region were conserved among all species. With the exception of an additional cysteine residue between positions 189 and 190 in the bovine sequence, the three other cysteine residues (amino acids 245, 253 and 269) found in the four SPARC sequences were also conserved, perhaps indicating a high degree of similarity in their tertiary structures. Also, on the basis of the consensus sequence Asn-Xaa-Ser/Thr, both mouse and *Xenopus* SPARC have a single *N*-glycosylation

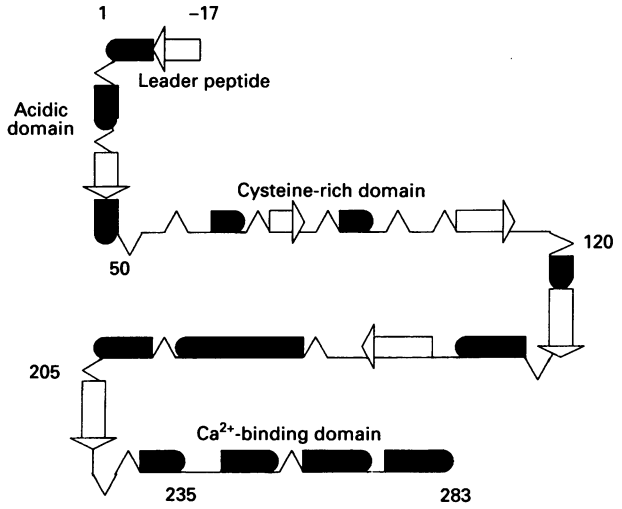

Fig. 4. Secondary structure of *Xenopus* SPARC

Putative secondary structure of *Xenopus* SPARC based on amino acid sequence predicted by the Chou–Fasman (1978) method. Bullet shapes (●) represent possible α-helical regions. Arrows (⇨) represent possible β-sheet regions. Acute angles (⩘) represent possible β-turns. Straight lines (—) represent possible random coil.

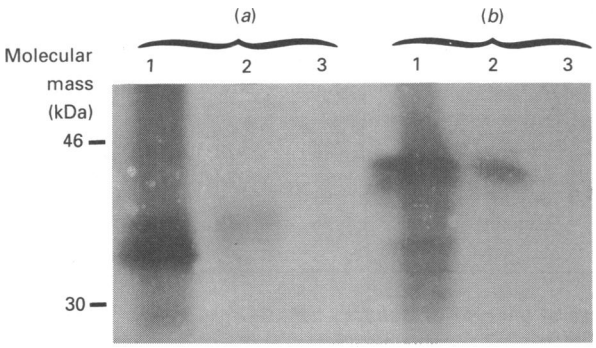

Fig. 5. Western-blot analysis of *Xenopus* SPARC

Protein samples were fractionated by SDS/PAGE on 12% gels, transferred to nitrocellulose, and probed with antisera to bovine SPARC. (*a*) Electrophoresis was performed under non-reducing and reducing (100 mM-dithiothreitol) conditions in (*b*). Samples in each lane were judged to be of roughly equal amounts by Coomassie Blue staining of a duplicate gel (results not shown). Duplicate samples were run in (*a*) and (*b*). Lane 1, approx. 1 μg of purified bovine osteonectin (generously provided by Dr. J. Sodek and Dr. C. Domenicucci, University of Toronto); lane 2, f.p.l.c.-purified *Xenopus* SPARC; lane 3, EDTA extract of total *Xenopus* long-bone proteins.

site at position 96. In comparison, bovine and human have two potential *N*-glycosylation sites at positions 68 and 96. Six possible phosphorylation sites exist in *Xenopus* SPARC, i.e. serine residues 84, 157, 188 and 228, and threonine residues 115 and 246, consistent with evidence indicating SPARC is phosphorylated.

On the basis of sequence similarity to the consensus Ca^{2+}-binding EF hands (Kretsinger, 1980), putative Ca^{2+}-binding sites at positions 160–173 and 255–265 in human SPARC have been assigned (Villareal *et al.*, 1989). However, secondary-structure analysis by the Chou–Fasman (1978) algorithm does not strongly support the presence of the classical helix–loop–helix structure of EF hands surrounding the putative Ca^{2+}-binding site at position 160. In contrast, such a helix–loop–helix structure seems possible at position 255 in *Xenopus*.

Size conservation between *Xenopus* and bovine SPARC

F.p.l.c.-fractionated *Xenopus* long-bone SPARC migrated at 39 kDa under non-reduced conditions on SDS/PAGE (Fig. 5*a*, lane 2), compared with 43 kDa under reduced conditions (Fig. 5*b*, lane 2). The doublet and shift is diagnostic of SPARC (Otsuka *et al.*, 1984). The *Xenopus* SPARC appeared similar in size to bovine SPARC (Figs. 5*a* and 5*b*, lane 1). No signal was observed when equivalent amounts of *Xenopus* long-bone EDTA extracts were probed (Figs. 5*a* and 5*b*, lane 3).

DISCUSSION

We present data demonstrating an equivalent of mammalian SPARC in *Xenopus*. On the basis of amino acid sequences derived from *Xenopus* and mammalian cDNAs, SPARC can be subdivided into three distinct highly conserved domains. The *N*-terminal glutamic acid-rich domain is the least conserved domain. Since this region can bind more than eight Ca^{2+} ions (Engel *et al.*, 1987) and has been speculated to be the hydroxyapatite-binding site (Bolander *et al.*, 1988), it may reflect differences in SPARC's abilities to interact with and/or inhibit hydroxyapatite-seeded crystal growth (Doi *et al.*, 1989). This could in part explain enrichment of SPARC in bovine, porcine and human osteoid tissues, in contrast with the low levels found in rodent (Zung *et al.*, 1986) and *Xenopus* osteoid tissues.

Another structural difference which may influence the distribution and function of SPARC within tissues is that both *Xenopus* and mouse have only one potential *N*-glycosylation site (Mason *et al.*, 1986), in contrast with the two putative sites in bovine (Bolander *et al.*, 1988) and human (Villareal *et al.*, 1989) SPARC. The high affinity of bone SPARC, relative to platelet SPARC, for collagen types I, III and V has been attributed primarily to differences in glycosylation (Kelm & Mann, 1991). However, Western analysis has not demonstrated significant M_r differences between bovine and *Xenopus* SPARC, despite different *N*-glycosylation potentials.

A single high-affinity EF-hand Ca^{2+}-binding site is present in all four species. The high affinity for Ca^{2+} ($K_d = 3 \times 10^{-7}$ M) demonstrated for this site in mouse (Sage *et al.*, 1989) indicates that it is fully occupied with Ca^{2+} at physiological concentrations. Hence, this site is not likely to be involved in Ca^{2+}-mediated regulatory events.

Several factors led to our examination of *Xenopus laevis* SPARC. A comparison of frog and mammalian SPARC amino acid sequences could shed light on important active regions of SPARC. *Xenopus laevis* also offers several biological and technical advantages, including a large egg, which can be easily manipulated, synchronous fertilization, and development, which can be easily followed *in vitro* (Sargent & Dawid, 1983). These advantages permit an examination of SPARC expression and function from fertilization to the adult state. We recently demonstrated the SPARC is expressed from the onset of neurulation in *Xenopus*. Interference with its expression leads to death of the embryo after neurulation (S. Damjanovski, F. Liu & M. Ringuette, unpublished work).

We thank Ms. Helen Heliopoulos and Dr. Jasmine Gruia-Gray for their assistance in preparation of the manuscript. We would like to extend special thanks to Dr. Jaro Sodek for providing us with materials and technical assistance. This work was supported by a National Science and Engineering Research Council grant to M.R.

REFERENCES

Bolander, M. E., Young, M. F., Fisher, L. W., Yamada, Y. & Termine, J. D. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 2919–2923

Chou, P. Y. & Fasman, G. D. (1978) Adv. Enzymol. Relat. Areas Mol. Biol. **47**, 45–148

Doi, Y., Okuda, R., Takezawa, Y., Shibata, S., Moriwaki, Y., Wakamatsu, N., Shimizu, N., Moriyama, K. & Shimokawa, H. (1989) Calcif. Tissue Int. **44**, 200–208
Domenicucci, C., Goldberg, H. A., Hofmann, T., Isenman, D., Wasi, S. & Sodek, J. (1988) Biochem. J. **253**, 139–151
Engel, J., Taylor, W., Paulsson, M., Sage, H. & Hogan, B. L. M. (1987) Biochemistry **26**, 6958–6965
Henikoff, S. (1987) Methods Enzymol. **155**, 156–165
Holland, P. W. H., Harper, S. J., McVey, J. H. & Hogan, B. L. M. (1987) J. Cell Biol. **105**, 473–482
Kelm, R. J., Jr. & Mann, K. G. (1990) Blood **75**, 1105–1113
Kelm, R. J., Jr. & Mann, K. G. (1991) J. Biol. Chem. **266**, 9632–9639
Kretsinger, R. H. (1980) CRC Crit. Rev. Biochem. **8**, 119–190
Kyte, J. & Doolittle, R. F. (1982) J. Mol. Biol. **157**, 105–132
Mason, I. J., Murphy, D., Munke, M., Francke, U., Elliott, R. W. & Hogan, B. L. M. (1986) EMBO J. **5**, 1831–1837
McDonald, J. A. (1988) Annu. Rev. Cell Biol. **4**, 183–207
Otsuka, K., Yao, K.-L., Wasi, S., Tung, P. S., Aubin, J. E., Sodek, J. & Termine, J. D. (1984) J. Biol. Chem. **259**, 9805–9812
Promega (1989) Promega Protocols and Application Guide, Promega Corporation Ltd., Madison
Ringuette, M. J., Damjanovski, S. & Wheeler, D. (1991) Biochem. Cell Biol. **69**, 245–250
Romberg, R. W., Werness, P. G., Riggs, B. L. & Mann, K. G. (1986) Biochemistry **25**, 1176–1180
Sage, H., Vernon, R. B., Funk, S. E., Everitt, E. A. & Angello, J. (1989) J. Cell Biol. **109**, 341–356
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
Sargent, T. D. & Dawid, I. B. (1983) Science **222**, 135–139
Termine, J. D., Kleinman, H. K., Whitson, S. W., Conn, K. M., McGarvey, M. L. & Martin, G. R. (1981) Cell (Cambridge, Mass.) **26**, 99–105
Towbin, H., Staehelin, T. & Gordon, J. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 4350–4354
Villareal, X. C., Mann, K. G. & Long, G. L. (1989) Biochemistry **28**, 6483–6491
Wasi, S., Otsuka, K., Yao, K.-L., Tung, P. S., Aubin, J. E., Sodek, J. & Termine, J. D. (1984) Can. J. Biochem. Cell Biol. **62**, 470–478
Zung, P., Domenicucci, C., Wasi, S., Kuwata, F. & Sodek, J. (1986) Biochem. Cell Biol. **64**, 356–362

Received 29 May 1991/16 July 1991; accepted 25 July 1991