Journal of Cellular Biochemistry 57:379–383 (1995)

PROSPECTS

Osteocalcin Cluster: Implications for Functional Studies

Christelle Desbois and Gerard Karsenty

Department of Molecular Genetics, The University of Texas, M.D. Anderson Cancer Center, Houston, Texas 77030

Abstract Osteocalcin is a skeletal member of the family of extracellular mineral binding Gla protein. Osteocalcin is synthesized only by the osteoblast and it is secreted into the bone matrix at the time of bone mineralization. The mineral binding properties of osteocalcin as well as its spatial and temporal pattern of expression suggest that it plays a role during bone mineralization, however until now its biological function is unclear. To understand osteocalcin function during skeletogenesis we mutated the two osteocalcin genes by homologous recombination in embryonic stem (ES) cells. Eight targeted clones were identified by Southern analysis using external probes. One of these clones contributed to the germ line of mouse chimera. Interbreeding of heterozygotes is currently in progress. Mutant mice will be useful to understand osteocalcin function in vivo. © 1995 Wiley-Liss, Inc.

Key words: bone extracellular matrix, cell proliferation, cell differentiation, osteoblast, odontoblast, osteocalcin, Gla protein

Osteoblasts or bone-forming cells are cells of mesodermal origin that produce the bone extracellular matrix (ECM) and the factors that mineralize it. This cell type has several interesting features from the point of view of cell biologists and molecular biologists. Once the osteoblasts are fully differentiated they exit the cell cycle and start producing bone ECM; in that respect these cells provide a very good model for studying the interrelationship between proliferation and differentiation. The second interesting feature of osteoblasts is that they are the only cells of the body that can mineralize the ECM they produce. Most likely there is a gene or genes specifically expressed in osteoblasts that fulfill this function.

Recently, very few genes have been shown to be expressed only in osteoblast and in odontoblast, its tooth counterpart. One of the better characterized osteoblast-specific genes is the gene coding for osteocalcin [1]. Osteocalcin is a member of a larger family of mineral-binding extracellular matrix proteins called Gla proteins. Until recently Gla proteins had been identified only in liver and skeleton. The liver Gla proteins include major coagulation factors, such as factor VII, IX, and X; prothrombin; and proteins S and C [2]. The functional importance of the Gla residues in these proteins is illustrated by clinical observations. A point mutation in the propeptide coding region of the factor IX gene leads to inhibition of γ-carboxylation of factor IX and eventually hemophilia β [3]. In another example, patients who are congenitally protein S deficient are predisposed to recurrent thrombosis [4]. It is interesting to note that two structurally homologous proteins have been adapted to serve both coagulant (i.e., factor IX) and anticoagulant (i.e., protein S) functions. To date two skeletal Gla proteins have been identified, sequenced, and subsequently cloned: osteocalcin and matrix Gla protein, which is present in cartilage as well as in bone.

Osteocalcin is one of the most abundant proteins synthesized by the osteoblast. As mentioned above, the protein is synthesized only by osteoblasts and odontoblasts *once they are fully differentiated.* Because of the existence of three Gla residues, osteocalcin has a high affinity for calcium and phosphate ions and hydroxyapatite crystals, and it is secreted by the osteoblasts into the bone ECM at the time of bone mineralization. Because of its restricted spatial and temporal pattern of expression and its structural features, osteocalcin has been suspected for a long time to play a role in bone mineralization, al-

Received August 5, 1994; accepted August 8, 1994.

Address reprint requests to Gerard Karsenty, Department of Molecular Genetics/Box 045, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030.

though its biological function is still unknown nearly 20 years after its original description. Recently, several groups have shown in several vertebrate species that there exists a cluster of osteocalcin and osteocalcin-related genes [5–7]. We have shown that these three genes are transcribed in different tissues during development and in different tissues in adult life, raising new questions about the function of the protein as well as the regulation of expression of these genes [7].

STRUCTURE OF THE MURINE OSTEOCALCIN GENE CLUSTER

To study the regulation of expression of osteocalcin gene and the function of the protein in an animal model convenient for genetic experimentation, we and others decided to clone the mouse gene [6,7]. When a mouse genomic library was screened with the mouse osteocalcin cDNA as a probe we observed hybridization to restriction fragments scattered over more than 15 kb of genomic DNA. This result was surprising since coding sequences of the human and rat osteocalcin gene are shorter than 1 kb. It suggested that several genes or pseudogenes with related sequences were in this genomic region. Indeed, Southern blot hybridization of genomic DNA cut with different restriction enzymes and hybridized with the mouse osteocalcin cDNA showed the existence of three bands. Through an extensive DNA sequencing effort we to showed that a) the three genes are side by side in the genome, and b) their coding sequence is highly similar (96% identity at the nucleotide level), and in particular the glutamic acid coding codons are all present in the three genes. The 5′ untranslated sequence of the genes located at the 5′ end of this cluster and of the gene located in the middle of the cluster were 93% identical over 1 kb and highly homologous to the 5′ untranslated sequence of the rat osteocalcin gene [8], so these two genes were called osteocalcin gene 1 (OG_1) and osteocalcin gene 2 (OG_2). The 5′ untranslated region of the gene located at the 3′ end of the cluster had a more complicated

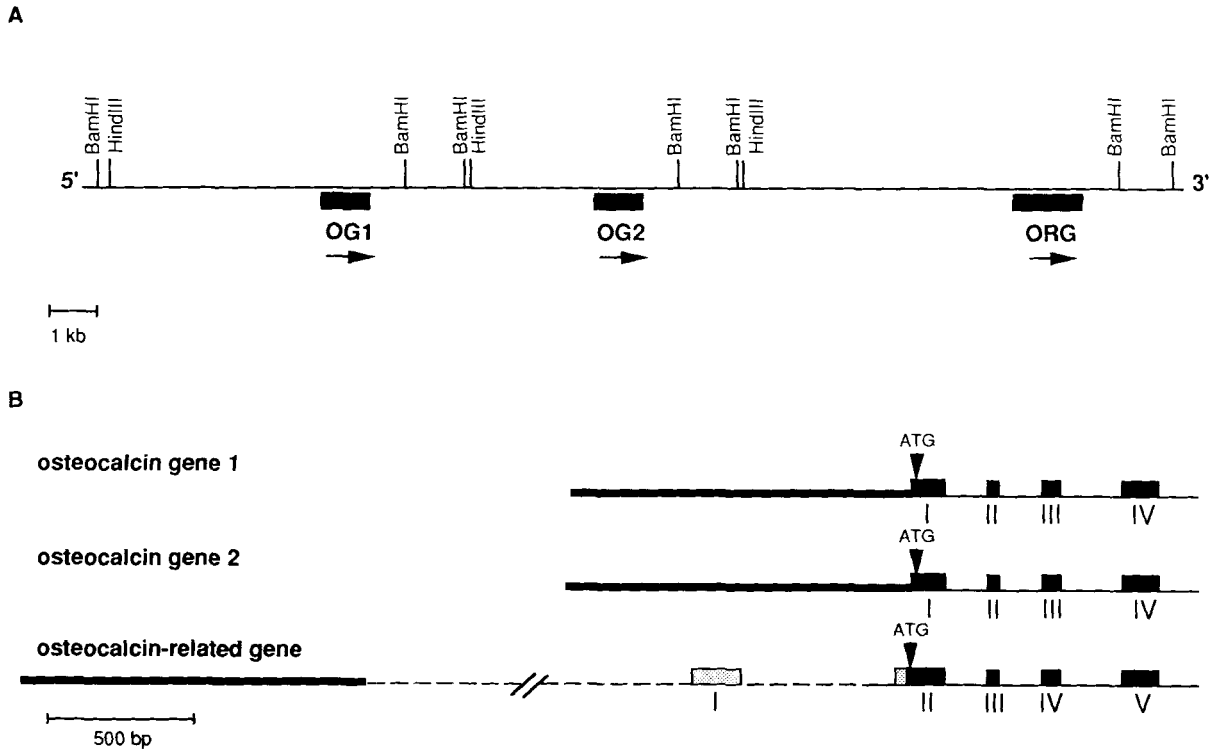

Fig. 1. Structure of the mouse osteocalcin gene cluster. **A:** A map of the 23-kb region containing the three genes is presented in relation to the phage clones used. Position of the three genes is shown by solid boxes. Arrows denote the orientation of transcription. Regions of the cluster that have been sequenced are depicted below by lines. **B:** Structural organization of each gene of the cluster. Solid boxes indicate common exons to the three genes, and shaded boxes depict exonic sequences specific to ORG. The solid lines 5′ of the ATG denote the 5′-untranslated sequences common to the three genes. The dashed lines 5′ of the exon 1 of ORG represent a 4-kb fragment specific of ORG.

organization. Upstream of the coding sequence there was a 4-kb DNA fragment that had no homology to any sequences in the two genes. This 4-kb sequence contained an intron, a noncoding exon, and a promoter. 5′ of the 4-kb insert there was a segment of DNA 93% homologous to the 5′ untranslated region of OG_1 and OG_2 over 1 kb. The existence of a different promoter suggested that this third gene was not expressed in the same tissues, so we called it osteocalcin related-gene (ORG).

The most intriguing aspect of this cluster of three genes is that each is expressed at different times during embryonic development and in different tissues in postnatal life. Using allele-specific hybridization of reverse transcriptase polymerase chain reaction products as depicted in Figure 2, we were able to show that OG_1 and OG_2 are expressed only in bone, whereas ORG is transcribed in kidney but not in bone. Furthermore, during embryogenesis OG_1 and OG_2 begin to be expressed at day 15.5 while ORG is transcribed starting at day 10.5. The level of expression of ORG is about 1% of the level of expression of OG_1 and OG_2.

In summary, the three genes of the cluster theoretically encode the same protein and are expressed at two major sites of calcium metabolism: bone, which is the main calcium ion reservoir in the body, and kidney, where calcium elimination occurs. The structure of this cluster and the spatial pattern of expression of the three genes raise several important questions: What is the function of Gla proteins in bone and in kidney? How is the distinct temporal and spatial pattern of expression achieved? Other clusters of developmentally regulated genes like the globin genes have been shown to share common regulatory elements that can dictate their sequential pattern of expression. It is conceivable that these 3 genes share common regulatory elements, which could turn on their expression sequentially. In the rest of this review we will describe the different approaches that have been used or could be used to study osteocalcin function.

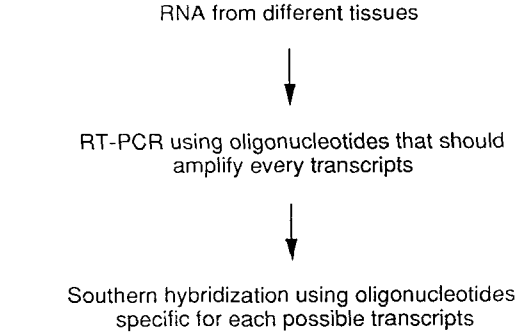

Fig. 2. Strategy to identify the different transcripts originating from the mouse osteocalcin gene cluster. An aliquot of each RTPCR reaction was electrophoresed on an agarose gel and subsequently hybridized with appropriate oligonucleotides. Plasmids containing each genomic clone were used as positive controls.

STUDY OF THE FUNCTION OF OSTEOCALCIN

Attempts to understand the function of osteocalcin have relied on in vitro experiments, a pharmacological model, and clinical investigations. In vitro, the transition from brushite ($CaHP^{42}H^{2}$) to hydroxyapatite [$Ca10(PO_4)^{6}(OH)^{2}$] is inhibited by osteocalcin, and this inhibitory function requires the presence of Gla residues [9]. These experiments suggest that osteocalcin acts as a negative regulator of bone mineralization, possibly preventing an overgrowth of the hydroxyapatite crystal. This would be consistent with its late temporal pattern of expression during bone formation. Further evidence supporting the hypothesis that osteocalcin acts as a negative regulator of bone mineralization arises from experiments in which bone particles were implanted subcutaneously to study bone resorption [10]. Bone particles were prepared from rats virtually depleted of osteocalcin by treatment with warfarin, an inhibitor of vitamin-K dependent carboxylase, for 6 weeks. Without the carboxylated residues, osteocalcin does not accumulate in bone matrix. These bone particles, devoid of osteocalcin, were resistant to resorption when implanted subcutaneously in normal rats. Likewise, these osteocalcin-deficient bone particles had a decreased ability to recruit and differentiate osteoclast progenitors compared to control bone particles. Taken together, these experiments suggest that osteocalcin may recruit cells of the osteoclast lineage to the site of newly formed bone, thereby regulating bone formation. Lastly, it has been shown in vitro that osteocalcin can form a complex with another noncollagenous protein of the bone ECM, osteopontin. Osteopontin is a cell adhesion protein that can interact with osteoclasts [11]. Conceivably an osteocalcin-osteopontin complex could recruit osteoclasts to a site of bone resorption and mediate the attachment of these cells at this particular site.

Attempts to study osteocalcin function in vivo have relied on the use of an anticoagulant drug called warfarin. Warfarin inhibits the vitamin-K-dependent synthesis of Gla residues in liver and bone proteins. It results in the intracellular accumulation of the precursors of osteocalcin and of matrix Gla protein that are not γ-carboxylated [12]. When pregnant rats were treated from gestational day 8 to 22 with a quantity of warfarin sufficient to halve the level of extracellular skeletal Gla protein 2, they showed minor modifications of the hypertrophic cartilage zone. However, when the level of skeletal Gla protein was decreased to 10% of control levels, a much more severe phenotype appeared, characterized by severe growth retardation in newborn rats and a decreased number of ossification centers. Histologically the endochondral growth plate was totally disorganized with an absence of alignment of hypertrophic chondrocytes and appearance of pericellular mineralization [13]. These findings in rat are similar to abnormalities observed in offspring of women to whom warfarin was administered during the first trimester of pregnancy. These malformations include nasal septum hypoplasia and stippling of epiphyses [14].

Although the in vivo experiments in rats using warfarin are informative, they must be interpreted cautiously. First, this treatment affects γ-carboxylation of glutamic acid residues present in osteocalcin, matrix Gla protein, and conceivably other yet to be characterized skeletal Gla proteins; therefore, the phenotype cannot be ascribed to a decreased level of a single protein. Second, these animals still had detectable levels of osteocalcin and in that regard should be considered hypomorphic rather than true "loss of function" animals. Nevertheless, based on these experiments we speculate that skeletal Gla proteins, and among them osteocalcin, play a role in the regulation of bone mineralization possibly by recruiting osteoclast to the site of newly formed bone and/or by preventing the growth of the hydroxyapatite crystal.

Measurement of serum osteocalcin, whose level reflects the portion of newly synthesized protein that does not bind to the mineral phase of bone, has been used as an indicator of bone formation in many physiological and pathological conditions. In general, serum osteocalcin concentrations correlate with histomorphometric indices of newly formed bone, leading most investigators to consider serum osteocalcin levels an indicator of bone formation. For example, serum osteocalcin concentration is increased in lactation, Paget's disease, and hyperparathyroidism and is decreased in pregnancy, estrogen or growth hormone deficiency, and alcoholism.

Many of the functions that have been attributed to osteocalcin are based solely on correlative arguments. Therefore, the availability of an experimental animal that lacks osteocalcin would be a great resource to study the role of osteocalcin during bone formation. Gene targeting and mouse embryonic stem (E.S.) cell technologies provide the ideal methodologies for generating such "loss of function" mutant animals [9,16,17]. These experiments would likely be more useful than overexpression of a mutated osteocalcin gene since all the mutations described in other Gla proteins so far are recessive.

E.S. cells are totipotent cells derived from mouse blastocysts. When reintroduced into a blastocyst, E.S. cells can contribute efficiently to the formation of all tissues in a chimeric mouse, including the germ line. Generation of mice in which a specific gene has been mutated involves the introduction of a mutation into a cloned genomic sequence of a particular locus. The mutation is then transferred by homologous recombination to the E.S. cell genome. Microinjection of the mutant E.S. cells into a mouse blastocyst should generate chimeras. These manipulated embryos are then transferred to the uterus of a pseudopregnant foster mother for development to term. If the E.S. cells have contributed to the germ line of the chimeras, some of the progeny of these chimeras will possess E.S. cell-derived genes, usually assessed initially by coat color markers. Thus, gene targeting E.S. cells provides a means for generating mice carrying specific mutations. Interbreeding of heterozygous mice yields animals homozygous for the mutation.

Considering the fact that the three genes of the mouse osteocalcin cluster are expressed in different tissues, two strategies can be proposed to achieve this gene deletion experiment. One approach would be to perform a deletion of the two genes expressed only in bone and a deletion of the gene expressed in kidney in separate experiments. The advantage of this strategy is that it would allow us to address questions about each gene in the cell type where it is expressed. Since ORG is expressed at very low levels compared to OG_1 and OG_2 and there is no evidence that the product of the gene can enter the gen-

eral circulation, the deletion of the two bone-specific genes should give valuable information on osteocalcin function during bone development. One possible pitfall that should be considered in this approach is that the level of expression of the ORG can be altered by the deletion of neighboring genomic sequences, but this can be easily monitored. Recent examples of gene deletion experiments have shown that it is hazardous to make any predictions of a phenotype [18]. However, the fact that osteocalcin does not start to be expressed before the late part of embryonic development suggests that osteocalcin-deficient mice should be alive at birth. The appearance of a phenotype at later stages will depend on whether other gene products could fulfill some or all of osteocalcins functions. In the case of deletion of ORG coding sequence it is even harder to predict a phenotype since we do not yet know the spatial pattern of expression of this gene and its gene product during embryonic development.

A second approach that is not necessarily exclusive of the first would be to perform a large deletion of the three genes of the cluster. The frequency of correct gene replacement could be much lower in deletion vectors bigger than 15 kb. The putative phenotype due to a deletion of the three genes could be difficult to interpret if it involves abnormalities in both kidney and skeleton, but it could also provide additional information.

SUMMARY

Osteocalcin is one of the most highly abundant proteins produced by osteoblasts and yet its biological function is still unknown. In the absence of any known genetic human disease linked to a mutation in one of the genes coding for osteocalcin, the best approach to characterizing osteocalcin functions in vivo is to delete the osteocalcin genes through homologous recombination in E.S. cells. This work will undoubtedly bring very important information about the molecular mechanisms of bone mineralization and may shed new light on the physiology of Gla proteins.

ACKNOWLEDGMENTS

This work was supported by NIH grant AR 41059-02, March of Dimes grant # 1FY91-0355, and a Roussel Institute Grant.

REFERENCES

1. Poser JW, Price PA (1979): J Biol Chem 254:431–436.
2. Furie B, Furie BC (1988): Cell 53:505–518.
3. Diuguid DL, Rabiet MJ, Furie BC, Liebman HA, Furie B (1989): Proc Natl Acad Sci USA 83:5803–5807.
4. Comp P, Nixon R, Cooper MR, Esmon, C (1984): J Clin Invest 74:2082–2088.
5. Rahman S, Oberdorf A, Montecino M, Tanhauser SM, Lian JB, Stein GS, Laipis PJ, Stein JL (1993): Endocrinology 133:3050–3053.
6. Desbois C, Hogue DA, Karsenty G (1994): J Biol Chem 269:1183–1190.
7. Sztainkrycer MD, Dow DA, Rowe DA, Pan LC (1993): J Bone Miner Res 8:5121.
8. Lian J, Stewart C, Puchacz E, Mackowiak S, Shalhoub V, Collart D, Zambetti G, Stein G (1989): Proc Natl Acad Sci USA 86:1143–1147.
9. Rombert RW, Werness PG, Riggs BL, Mann KG (1986): Biochemistry 25:1176–1180.
10. Lian J, Tassinari M, Glowacki J (1984): J Clin Invest 73:1223–1226.
11. Ritter N, Farach-Carson M, Butler W (1992): J Bone Miner Res 7:877–885.
12. Suttie JW (1985): Annu Rev Biochem 56:459–477.
13. Feteih R, Tassinari MS, Lian JS (1990): J Bone Miner Res 5:885–894.
14. Pauli R, Lian JS, Mosher D, Suttie JW (1987): Am J Hum Genet 41:566–583.
15. Hauschka PV, Lian JB, Cole DEC, Gundberg CM (1989): Physiol Res 69:990–1047.
16. Capechi MR (1989): Trends Genet 5:70–76.
17. Robertson EJ (1986): Trends Genet 2:9–13.
18. Soriano P, Montgomery C, Geske R, Bradley A (1991): Cell 64:693–702.