CREB-AP1 Protein Complexes Regulate Transcription of the Collagen XXIV Gene (Col24a1) in Osteoblasts*

Noritaka Matsuo†§, Shizuko Tanaka†, Marion K. Gordon‡, Manuel Koch¶, Hidekatsu Yoshioka‡, and Francesco Ramirez**††

From the †Laboratory of Genetics and Organogenesis, Research Division of the Hospital for Special Surgery, and Department of Physiology and Biophysics, Weill Medical College, Cornell University, New York, New York 10021, the ‡Department of Anatomy, Biology, and Medicine, Oita University, I-1 Hasama-machi, Oita 879-5593, Japan, the §Enviromental and Occupational Health Sciences Institute and the Department of Pharmacology and Toxicology, Ernest Mario School of Pharmacy, Rutgers University, Piscataway, New Jersey 08854, the ¶Institute for Biochemistry II, University of Cologne, 50931 Cologne, Germany, and **CEINGE-Biotecnologie Avanzate, 80131 Naples, Italy

Collagen XXIV is a newly discovered and poorly characterized member of the fibril-forming family of collagen molecules, which displays unique structural features of invertebrate fibrillar collagens and is expressed predominantly in bone tissue. Here we report the characterization of the proximal promoter of the mouse gene (Col24a1) and its regulation in osteoblastic cells. Using well characterized murine models of osteoblast differentiation, we found that the Col24a1 gene is activated sometime before onset of the late differentiation marker osteocalcin. Additional analyses revealed that Col24a1 produces equal amounts of two alternatively spliced products with different 5′-untranslated sequences that originate from distinct transcriptional start sites. Cell transfection experiments in combination with DNA binding assays demonstrated that Col24a1 promoter activity in ROS17/2.8 osteosarcoma cells is under the control of an upstream cis-acting element, which is shared by both transcripts and is recognized by specific combinations of c-Jun, CREB1, ATF1, and ATF2 dimers. Consistent with these results, overexpression of c-Jun, ATF1, ATF2, or CREB1 in transiently transfected osteoblastic cells stimulated transcription from reporter gene constructs driven by the Col24a1 promoter to different degrees. Moreover, chromatin immunoprecipitation experiments showed that these nuclear factors bind the same upstream sequence of the endogenous Col24a1 gene. Collectively these data provide new information about transcriptional control of collagen fibrillogenesis, in addition to implicating for the first time CREB-AP1 protein complexes in the regulation of collagen gene expression in osteoblasts.

Vertebrate collagens represent a very large superfamly of extracellular proteins that impart specific physical properties to the connective tissue of virtually every organ system (1–3). There are more than 42 collagen α-chains that form 27 distinct trimers or types, which in turn give rise to a large variety of specialized macroaggregates. The most abundant and ubiquitous collagen macroaggregates are the highly ordered banded fibrils made of the so-called fibrillar collagens (types I-III, V, and XI) (1–3). All members of the fibrillar collagen family share a common structure that consists of a long triple helical domain, which is made of uninterrupted Gly-X-Y triplets and flanked at both ends by noncollagenous propeptides (1–3). These structural features also characterize collagen molecules that form fibrils in the extracellular matrices of primitive invertebrates, such as sponges, annelids, echinoderms, and mollusks (4, 5). Unlike the vertebrate counterparts, invertebrate fibrillar collagens display short interruptions in the triple helices and unique structural features in the amino- and carboxyl-terminal propeptides (4, 5). Vertebrate fibrillar collagens are either widely distributed in soft and hard tissues (types I, III, and V) or are restricted predominantly to cartilage (types II and XI) (1). Genetic evidence from animal and human studies has indicated that the quantitatively minor types V and XI collagen regulate the diameter of the major types I and II fibrils, respectively, by participating in fibril assembly (2, 6–9). These studies have also demonstrated the importance of heterotypic collagens I/V and II/VI fibrils in skeletal development and integrity.

As a result of the Human Genome effort, several new collagens have been recently identified which had escaped prior biochemical detection. Two among them (types XXIV and XXVII) bear the structural characteristics of fibrillar collagens and specifically, of invertebrate fibrillar collagens (10–12). Gene expression analyses in the mouse have revealed that Col24a1 and Col27a1 display mutually exclusive patterns in the developing and adult skeleton. These studies have in fact shown that whereas Col24a1 transcripts accumulate at ossification centers of the craniofacial, axial, and appendicular skeleton, Col27a1 activity is instead confined to the cartilaginous anlagen of skeletal elements (10–12). Additionally, structural considerations have suggested that collagens XXIV and XXVII are likely to form distinct homotrimers (11). Together these observations have been interpreted to indicate that these newly discovered fibrillar collagens may participate in the control of important physiological processes in bone and cartilage, such as collagen fibrillogenesis and/or matrix calcification and mineralization (10–12).

Bone formation is a complex and tightly regulated genetic program that involves two distinct pathways at different anatomical locations (13–15). In intramembranous ossification, mesenchymal cells condense and differentiate directly into collagen I-producing osteoblasts. In endochondral bone formation, cells at condensation sites differentiate into chondrocytes that form a cartilage (collagen II-rich) anlagen, which is replaced by a bony (collagen I-rich) matrix and bone marrow following chondrocyte hypertrophy, matrix calcification, and vascular invasion. At the same time, cells around the condensations form the perichondrial layer that gives rise to the osteoblast-forming perieosteum and ulti-
Collagen XXIV Regulation in Osteoblasts

Collagen XXIV is a type of collagen that plays a role in the extracellular matrix of bone and cartilage. It is particularly important in the regulation of bone homeostasis and the repair of cartilage. This study aimed to investigate the regulatory mechanisms of collagen XXIV expression in osteoblasts.

**EXPERIMENTAL PROCEDURES**

**Cell Cultures**—Rat osteosarcoma ROS17/2.8 and ROS25 cells, mouse fibroblast NIH-3T3, and mouse primary calvarial osteoblast cells (MCC) were used in this study. ROS17/2.8, ROS25, and NIH-3T3 were cultured in Dulbecco’s modified Eagle’s medium supplemented with 10% heat-inactivated fetal bovine serum at 37 °C in humidified 8% CO2 and 92% air (32). ROS17/2.8, ROS25 and NIH-3T3 were formed using a commercial kit (Upstate Biotechnology) (34). Quantitative PCR was carried out for 35 cycles using 5 μl of sample DNA solution/50-μl reaction, and amplification products were separated in 2.5% agarose gel in 1× TAE buffer.

**RESULTS**

Col24a1 Contains Two Alternative Promoters—Previous work established the entire coding sequence of the human α1(XXIV) collagen (COL24A1) gene and of only part of the mouse Col24a1 gene (11). We used this information to complete the primary structure of the mouse α1(XXIV) collagen chain by identifying mouse expressed sequence tags in the GenBank™ and by generating PCR amplification products covering sequence gaps. As a result, we found a 19-amino acid insertion in the noncollagenous amino-terminal sequence of the mouse compared with the human chain (see GenBank™ accession numbers AY244357 and DQ157748). These experiments were also instrumental in identifying the foremost exon of Col24a1 as consisting of a 5′-untranslated region (UTR) of undetermined length and a coding segment corresponding to the first 94 amino acid residues of the α-chain. The oligonucleotide-capping RACE approach was therefore employed to determine the Col24a1 start site of transcription and, implicitly, the

---

2 The abbreviations used are: CREB, cyclic AMP-responsive element-binding protein; AP1, activator protein 1; ATF, activating transcription factor; Chip, chromatin immuno precipitation; CRE, cyclic AMP-responsive element; EMSA, electrophoretic mobility shift assay; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; LUC, luciferase; MCC, mouse primary calvarial osteoblast cells; RACE, rapid amplification of cDNA ends; UTR, untranslated region.
5’-boundary of exon 1. As the source of template for the reaction, we utilized RNA purified from the eye and bone in which accumulation of Col24a1 transcripts has been found to be the highest (11). Sequencing of nearly 40 independent cDNA clones from each set of RNA samples revealed the presence of two different 5’-UTRs that upon subsequent analysis of genomic clones, were accounted for by a combination of alternative splicing and transcriptional start sites. To be precise, half of the cDNA clones contained a 353-nucleotide long 5’-UTR that is continuous with the genomic sequence of the exon originally identified as the first of exon of the human gene and which was now renamed exon 1a (Fig. 1A) (11). The other half of the cDNA clones contained a 353-nucleotide long 5’-UTR that is continuous with the genomic sequence of the exon originally identified as the first of exon of the human gene and which was now renamed exon 1b (Fig. 1A) (11). The other half of the cDNA clones instead contained the 87 nucleotides immediately upstream of the ATG codon, in addition to an 80-nucleotide long 5’-UTR corresponding to an upstream exon (named exon 1b) that is separated from exon 1a by a 152-bp intervening sequence (Fig. 1A).

Both transcripts 1a and 1b are spliced correctly into exon 2, leaving the open reading frame unaffected, and consequently, they are predicted to translate into identical 1(XXIV) chains (Fig. 1A).

In summary, Col24a1 contains two alternative start sites of transcription, thereby identified as –1 (exon 1a) and +232 (exon 1b), two alternatively spliced transcripts with different 5’-UTRs, the shortest of which (transcript 1b) splices into nucleotide +509 of exon 1a, and the same start site of translation, located at nucleotide +586 (Fig. 1A). The functional significance of Col24a1 alternative promoters and 5’-UTR heterogeneity was not addressed in the present study. Comparison of the 5’-end sequences of the COL24A1 and Col24a1 genes revealed three segments of high homology which span from nucleotides –100 to +133, from +198 to +256, and from +470 to +632 (Fig. 1A).

Collagen XXIV has been estimated to represent about 4% of the amount of collagen I in bone, thus slightly less than collagen V, the regulator of collagen I fibrillogenesis (11). The estimate was based on gene expression analyses that in addition documented coexpression of collagens I and XXIV genes at ossification centers in the mouse embryo (11). Osteoblasts were therefore chosen as the experimental system in which to study the transcriptional regulation of Col24a1. Reverse transcription-PCR amplifications were used to assess the levels of Col24a1 expression in ROS17/2.8 and ROS25 osteosarcoma cell lines, which represent late and early stages of osteoblast differentiation, respectively (33). As positive and negative controls, PCR amplifications were also performed with RNAs purified from MCC and NIH-3T3 fibroblasts. Osteoblast-specific genes included Col1a2 (early osteoblast differentiation marker) and osteocalcin (late osteoblast differentiation marker), whereas the ubiquitous GAPDH gene served as the normalizing control. The results of these experiments suggested that the onset of Col24a1 expression occurs sometime after Col2a1 and prior to osteocalcin gene activation (Fig. 1B). Implicitly, they also identified ROS17/2.8 cells as the most suitable model in which to study the anatomy of the minimal Col24a1 promoter.

**A Short Upstream Sequence Promotes Col24a1 Transcription in Osteoblasts**—Cell transfection experiments were initially employed to delineate the shortest promoter sequence of Col24a1 capable of direct-
ing transcription in ROS17/2.8 cells. To this end, we engineered two distinct sets of LUC reporter gene constructs representative of the alternative promoters of \textit{Col24a1}. The first set of \textit{Col24a1} promoter-LUC constructs shared the same 3' end at position +509 and included both start sites of transcription, whereas the 3' ends of the second set of \textit{Col24a1} promoter-LUC constructs was located at +80 and excluded the start site of transcript 1b (Fig. 2A). Both sets of promoter-LUC constructs included progressive 5'-deletions of the upstream \textit{Col24a1} sequence (Fig. 2A). Irrespective of the 3' end of the promoter-LUC construct, cell transfection assays assigned maximal transcriptional activity to the region between nucleotides -144 and +80 (Fig. 2B). This promoter segment contains one of the three homology sequences of the \textit{COL24A1} and \textit{Col24a1} genes (Fig. 1A).

Next, an electrophoretic mobility shift assay (EMSA) was employed to identify potential DNA-nuclear protein interactions within the -144 to +81 segment of the \textit{Col24a1} promoter. To this end, four overlapping probes (p1–p4) spanning from nucleotide -163 to nucleotide +116 were each incubated with ROS17/2.8 nuclear extracts (Fig. 3A). Specific band shifts were only obtained with overlapping probes p2 and p3, which cover together the sequence between nucleotides -98 and +51 (Fig. 3A). To be precise, p2 yielded four retarded bands (b1–b4) of which p3 appeared to migrate as band b2 of probe p3 (Fig. 3A). In support of this postulate, band b2 disappeared from the p2 EMSA pattern when competed with a molar excess of unlabeled probe p3; conversely, formation of the p3 retarded band was eliminated by competition with a molar excess of the p2 sequence (Fig. 3B). Taken together, these results mapped the b1, b3, and b4 binding sites between nucleotides -98 and -33 and the binding site of b2 between nucleotides -33 and -15 (Fig. 4).

The functional contribution of the segment encompassing the b1–b4 binding sites was evaluated by cell transfection experiments using the -144 + 509/LUC plasmid bearing internal deletions of the -98 to -33 sequence (p2D; b1,3,4 binding sites), the overlapping -33 to -15 sequence (p2/3D; b2 binding site), or both of them (p2,3D; b1–b4 binding sites) (Fig. 4A). Unlike elimination of the b2 binding site in the p2/3D construct, deletion of the upstream sequence that gives rise to retarded bands b1, b3, and b4 led to a drop in luciferase activity of construct p2D nearly equal to the -144 + 509/LUC construct with the internal dele-
element appears to drive most of common between probes p2 and p3. EMSAs demonstrating that binding of the b2 retarded complex is in the sequence in retarded complexes b1–b4 are shown on the nuclear extracts (Fig. 4). Inhibition of all binding sites (p2,3D) or the shorter –509/LUC plasmid (Fig. 4B). We therefore focused on the characterization of the factors binding to the sequence between –98 and –33 because this cis-acting element appears to drive most of Col24a1 promoter expression in ROS17/2.8 cells.

Binding of CREB-ATF and AP1 Complexes Is Required for Col24a1 Transcription—A series of shorter oligonucleotides encompassing the –98 to –33 sequence was used in the EMSA to compete in vitro binding to the p2 probe to narrow down the site(s) of nuclear protein interaction within this cis-acting element. The results clearly demonstrated that all p2 retarded complexes (b1, b3, and b4) were competed by the p2b oligonucleotide that extends from nucleotides –55 to –38 (Fig. 5A). Inspection of the –55 to –38 sequence identified within it the TGACGTCA sequence, which represents a perfect match of the cyclic AMP-responsive element (CRE) protein/ATF binding site (Fig. 5B) (36). Indeed competition experiments using a molar excess of a mutated CRE oligonucleotide failed to interfere with formation of the b1, b3, and b4 complexes (Fig. 5B). The significance of the binding assays was corroborated further by cell transfection experiments that showed a loss of promoter activity in the –144 + 509/LUC plasmids bearing the same p2b mutation (Fig. 4).

Members of the CREB-ATF family of nuclear factors bind with high affinity to CRE, but AP1 dimers can also interact with CREB-ATF sites, depending on the composition of their flanking sequences (36). Antibodies specific for AP1 and CREB family members were therefore used in a screen to identify which of the possible protein combinations bind the p2b element of Col24a1. The results showed supershifts or binding interferences only with c-Jun, CREB1, ATF1, and ATF2 antisera (Fig. 6).

Expression vectors for c-Jun, ATF1, ATF2, or CREB1 were each cotransfected into ROS17/2.8 cells together with the wild-type or p2a mutant –144 + 509/LUC plasmid to correlate DNA-protein binding with promoter function. These functional assays demonstrated that each of the four recombiantly expressed nuclear factors was capable of stimulating Col24a1 promoter activity in a dose-dependent manner and only when the integrity of the p2b sequence was preserved (Fig. 7A). Independent confirmation of the functional assays was obtained by a ChIP assay that documented in vivo occupancy by CREB1, ATF1, c-Jun, and ATF2 of the p2b element in the endogenous Col24a1 promoter (Fig. 7B). Consistent with the EMSA data, this in vivo assay also provided evidence for CREB-AP1 specificity by showing lack of JunD and ATF4 binding to the collagen but not the TIMP1 or osteocalcin promoters (Fig. 7B) (37, 38).

**DISCUSSION**

The present study demonstrated that Col24a1 is a marker of late osteoblast differentiation which is positively regulated by the binding of specific combinations of CREB-AP1 proteins to an upstream cis-acting element, which is shared by the two alternative promoters of the gene. These findings advance knowledge of the transcriptional pathways that regulate formation of fibrillar collagen assemblies in the skeleton. They will also inform the characterization of genetic programs that may be negatively affected by the loss of collagen XXIV in the developing and adult mouse bone.

The role of collagen I fibrils in bone physiology is well established as is the contribution of the minor collagen V to guiding nucleation of collagen I fibrils in the skin and the eyes and by extrapolation, in bone tissue (1–3, 6, 8, 9). The expression pattern of Col24a1 in the developing bone and the eye is at least consistent with a similar role of this macromolecule in collagen I fibrillogenesis, even though the lack of suitable antibodies has hampered experimental confirmation of this hypothesis. Our preliminary findings indicate that Col24a1 is inactive in ROS25 cells, which correspond to early differentiating osteoblasts, but not in ROS17/2.8 cells, which represent late differentiating osteoblasts, or in calvarial...
osteoblasts, albeit at seemingly lower levels than in ROS17/2.8 cells. Within the limitations of this experimental system, these results nonetheless suggest that collagen XXIV is an integral part of the genetic program of osteoblast terminal differentiation (39). That collagen XXIV is expressed at a lower level in nonskeletal tissues, such as the brain and the eye, also suggests a potentially broader role in organogenesis (11).

Our study adds AP1 and CREB-ATF proteins to the list of transcription factors that are involved in the regulation of fibrillar collagen genes, particularly in osteoblastic cells. A large body of work has demonstrated the critical contributions of these two families of basic leucine zipper proteins to bone formation and remodeling (40–42). For example, genetic alterations in functions of AP1 and related proteins have been shown to affect negatively osteoblast differentiation and function as well as bone development (40). Similarly, transgenic interference of CREB protein activity greatly impairs the normal process of endochondral bone formation (43). Members of the AP1 or CREB-ATF family of nuclear factors can form homodimeric or heterodimeric protein complexes, which transduce distinct signals and exert discrete transcriptional responses on various promoter targets (36). The heterogeneity in dimer composition is the main determinant of the functional diversification of AP1 and CREB-ATF complexes, which include dimers within and between selected members of each family of transcription factors (36, 40, 41). Obligatory combinations of CREB-ATF proteins recognize the octameric TGACGTCA element, as do heterodimers between ATFs and specific Jun and/or Fos proteins (36). In line with this last consideration, our antibody interference experiments indicate that CREB1, CREB1/ATF1 and c-Jun/ATF2 dimers can specifically bind in vitro to the evolutionarily retained TGACGTCA sequence (p2b element) of the Col24a1 and COL24A1 promoters. Further confirmation of this in vitro finding was provided by the ChIP assay, which showed that the endogenous p2b site of the rat gene is occupied in ROS17/2.8 nuclear extracts in the absence (−) or in the presence (+) of a 100-fold molar excess of unlabeled oligonucleotides p2a–c that cover the p2 sequence (nucleotide positions shown above the autoradiographs). B, top, the CREB-ATF binding site within the p2b sequence is highlighted, and the nucleotide substitutions (mt) are indicated (p2bmt, see also relevant functional assay of Fig. 4); bottom, EMSAs showing formation of retarded complexes between the p2b probe and ROS17/2.8 nuclear extracts in the absence (−) or in the presence (+) of increasing (10–100-fold) molar excess of unlabeled wild-type (wt) or mutant (mt) p2b oligonucleotides.
indicate that binding of these nuclear proteins stimulates transcription from the minimal Col24a1 promoter in osteoblasts. This conclusion is based on the absolute requirement of element 2b integrity for promoter activity and on the positive effect on promoter activity of each of the four nuclear proteins overexpressed in ROS17/2.8 cells. The expression profiles of the proteins that bind element 2b during osteoblast differentiation in vitro are also consistent with the time of Col24a1 onset estimated by the present study. Jun proteins are in fact highly expressed in differentiating osteoblasts prior to matrix production and mineralization, and phosphorylated CREB reaches its highest level in the early mineralization stage (44, 45). Irrespective of whether or not these findings are functionally correlated, the characterization of the Col24a1 promoter further support the emerging notion that distinct regulatory pathways coordinate expression of different fibrillar collagen genes in bone tissue. Transgenic studies have in fact indicated that cooperation between an Sp1 binding site in the proximal promoter and uncharacterized complex(es) interacting with a far upstream enhancer directs bone-specific expression of the human COLIA2 gene (46). The same kind of approach has identified positive osteoblast-specific elements in the upstream promoter of the mouse and rat Col1a1 genes, as well as a δEF1/ZEβ1 binding site further upstream which represses transcription of the mouse promoter in osteoblasts (47–49). Work in progress is extending the present work to the characterization of possible interactions of the CREB-AP1 complexes with other nuclear proteins, in addition to evaluating these in vitro findings within the physiological context of the transgenic mouse model.

Acknowledgments—We thank Bayan Sudan for the excellent technical assistance and Karen Johnson for organizing the manuscript. We are also indebted to Lionel Ivashkiv and Gerard Karsenty for providing critical reagents.

REFERENCES

Collagen XXIV Regulation in Osteoblasts


