ARTICLES

Characterization of the Upstream Mouse *Cbfa1/Runx2* Promoter

Z. S. Xiao, Shi-Guang Liu, T. K. Hinson, and L. D. Quarles*

Department of Medicine, Duke University Medical Center, Durham, North Carolina

Abstract   *Cbfa1 (or Runx2*/AML-3/*PEPB2α*) is a transcriptional activator of osteoblastic differentiation. To investigate the regulation of *Cbfa1* expression, we isolated and characterized a portion of the 5'-flanking region of the *Cbfa1* gene containing its ''bone-related'' or P1 promoter and exon 1. We identified additional coding sequence in exon 1 and splice donor sites that potentially give rise to a novel *Cbfa1* isoform containing an 18 amino acid insert. In addition, primer extension mapping identified in the *Cbfa1* promoter a minor mRNA start site located ∼0.8 kb 5' upstream of the ATG encoding the MASN/p57 isoform and ∼0.4 kb upstream of the previously reported start site. A luciferase reporter construct containing 1.4 kb of the mouse *Cbfa1* promoter was analyzed in Ros 17/2.8 and MC3T3-E1 osteoblast cell lines that express high levels of *Cbfa1* transcripts. The activity of this construct was also examined in non-osteoblastic Cos-7 and NIH3T3 cells that do not express *Cbfa1* and mesenchymal-derived cell lines, including CH3T101/2, C2C12, and L929 cells, that express low levels of mature *Cbfa1* transcripts. The 1.4 kb 5' flanking sequence of the *Cbfa1* gene directed high levels of transcriptional activity in Ros 17/2.8 and MC3T3-E1 osteoblasts compared to non-osteoblasts Cos-7 cells, but this construct also exhibited high levels of expression in C310T1/2, L929, and C2C12 cells as well as NIH3T3 cells. In addition, *Cbfa1* mRNA expression, but not the activity of the *Cbfa1* promoter, was upregulated in a dose-dependent manner in pluripotent mesenchymal C2C12 by bone morphogenetic protein-2 (BMP-2). These data indicate that *Cbfa1* is expressed in osteogenic as well as non-osteogenic cells and that the regulation of *Cbfa1* expression is complex, possibly involving both transcriptional and post-transcriptional mechanisms. Additional studies are needed to further characterize important regulatory elements and to identify additional regions of the promoter and/or post-transcriptional events responsible for the cell-type restricted regulation of *Cbfa1* expression. J. Cell. Biochem. 82: 647–659, 2001.   © 2001 Wiley-Liss, Inc.

Key words: transcription factors; *Cbfa1*; *Osf2*; osteoblasts; Runx2

The *Cbfa1* (core-binding factor) gene (also known as Runx2/AML-3/*PEPB2α*) encodes a transcription factor which plays a critical role in osteoblast-mediated bone formation [Ducy et al., 1997; Komori et al., 1997; Mundlos et al., 1997; Otto et al., 1997]. The role of *Cbfa1* gene products is documented by the failure of bone to form following targeted disruption of the *Cbfa1* gene in mice [Komori et al., 1997; Otto et al., 1997]. In addition, inactivating mutations of human CBFA1 cause cleidocranial dysplasia (CCD), an autosomal-dominant condition characterized by abnormal skeletogenesis and the arrest of osteoblast development as well as alterations in chondrocyte development [Mundlos et al., 1997]. Recent studies indicate that *Cbfa1* also regulates osteoblast function postnatally [Ducy et al., 1999].

The *Cbfa1* gene comprises at least eight exons [Mundlos et al., 1997; Geoffroy et al., 1998; Xiaoet al., 1998], referred herein as exons 1 through 8. The mouse gene potentially gives rise to several isoforms through alternative mRNA splicing, utilization of different translational start-codons and two separate promoters

designated P1 and P2 [Ogawa et al., 1993; Satake et al., 1995; Stewart et al., 1997; Geoffroy et al., 1998; Thirunavukkarasu et al., 1998; Xiao et al., 1998; Drissi et al., 2000]. Some of these isoforms have unique N-terminal sequences that may be required for optimal transactivation [Thirunavukkarasu et al., 1998]. The originally described *Cbfa1* gene product (which we will refer to as *Cbfa1 / p56* isoform or PEBP2αA) is expressed from the P2 promoter and utilizes a translation start site in exon 2 (previously referred to as exon 1, Fig. 1) and is predicted to be a 513-amino acid protein that begins with the sequence MRIPVD [Ogawa et al., 1993; Satake et al., 1995; Mundlos et al., 1997]. This isoform, a T-cell gene product [Satake et al., 1995], also has been detected in osteoblasts by some [Drissi et al., 2000; Harada et al., 1999] but not by other investigators [Ducy et al., 1997]. More recently, distal exons of the *Cbfa1* gene have been identified that potentially encode the N-termini of *Cbfa1* isoforms expressed in osteoblasts [Geoffroy et al., 1998; Thirunavukkarasu et al., 1998; Xiao et al., 1998]. These upstream exons contain the 5′ untranslated region (5′UTR) and the N-terminus of osteoblast specific *Cbfa1* transcripts (designated as *Cbfa1 / iso*, also referred to as *Cbfa1*/p57 and OSF2, Fig. 1) that encode a gene product beginning with the amino acid

sequence "MASNSL" [Stewart et al., 1997; Xiao et al., 1998]. This isoform is expressed from the P1 or "bone-related" upstream promoter [Drissi et al., 2000]. RNA splicing events generate additional *Cbfa1* isoform diversity [Geoffroy et al., 1998; Xiao et al., 1998]. Alternative 5′ splice site selection arising from the differential use of two splice donor sites in exon 1 potentially gives rise to *Cbfa1* transcripts (designated *Cbfa1*d1 and *Cbfa1*d2) that differ by 33 nucleotides in their 5′UTR [Xiao et al., 1998]. In addition, the fully processed *Cbfa1* mRNA requires a micro-splicing event to remove a mini-intron in exon 1 [Xiao et al., 1998]. Additional alternative splicing events have been identified in both mouse *Cbfa1* and human CBFA1 leading to the exclusion of various exons and mRNA splicing variants producing truncated proteins [Ogawa et al., 1993; Satake et al., 1995; Geoffroy et al., 1998; Xiao et al., 1998]. The presence of splicing variants encoding distinct *Cbfa1* proteins [Ogawa et al., 1993; Satake et al., 1995; Geoffroy et al., 1998; Xiao et al., 1998] suggests that pre-mRNA splicing may play a role in regulating cell-type specific *Cbfa1* expression.

To date, limited information is available regarding the mechanism of tissue and cell-type restricted transcriptional control of *Cbfa1* gene expression. To examine potential mechanisms
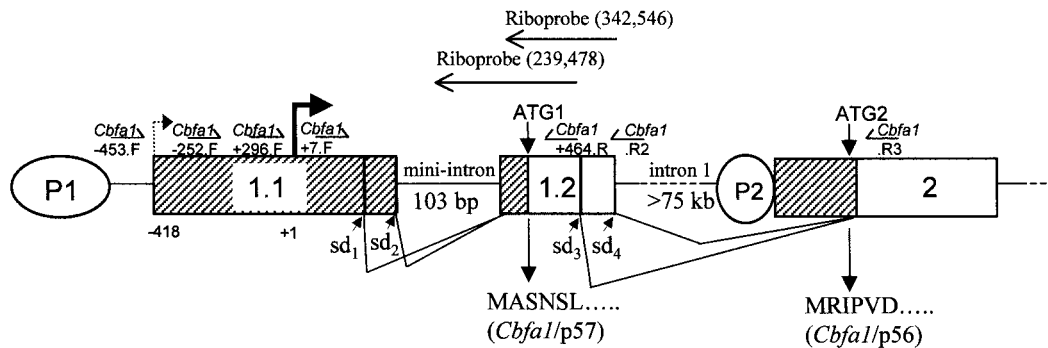
Fig. 1. Organization of the 5′ end of the *Cbfa1* gene. Exon 1 encodes the 5′ UTR (hatched area) and the unique N-terminal *Cbfa1* isoforms. Exon 1 is divided into two parts (designated 1.1 and 1.2 by a "mini-intron" in the 5′UTR of mRNAs CBfa1. Exons 1 and 2 each contain a single functional ATG translation start codon (ATG1 and ATG2) for the respective *Cbfa1*/p57 (or *Cbfa1/iso*) beginning with N-terminus MASNSL and *Cbfa1/p56* (or *Cbfa1/org*) beginning with the N-terminus MRIPVD. Exons 1 and 2 are separated by a large intron (intron 1). The sequence of the upstream P1 promoter (circle) is shown in Figure 2. Exon 1 has two splice donor sites (sd1 and sd2) in the 5′UTR and two additional splice donor sites (sd3 and sd4) in the coding region of exon 1. Alternative splice site selection generates additional transcripts, *Cbfa1*d1 and *Cbfa1*d2, that differ by 33 nucleotides in the 5′UTR of *Cbfa1*/p57 [Xiao et al., 1998] as well as a novel *Cbfa1* isoform containing an 18 amino acids insert, GKCY-LISSLFLCEPVKHE. The location of the nucleotide primers used for RT-PCR amplification of mRNAs are depicted by the arrows. The regions recognized by the riboprobes are also depicted. The putative minor transcription start site as defined by primer extension (see Fig. 3 below) is designated as −418 and the originally described start site as +1 nucleotide [Geoffroy et al., 1998].

for osteoblast-specific regulation of *Cbfa1* gene expression, we isolated the "bone-related" or P1 promoter of the *Cbfa1* gene and performed functional analysis using a promoter-reporter gene construct in osteoblastic and non-osteoblastic cell lines. Comparison of promoter activity and endogenous *Cbfa1* gene expression revealed that the P1 promoter only partially accounts for the cell-restrict expression of *Cbfa1*.

MATERIALS AND METHODS

Genomic Cloning

A mouse 129 SVJ genomic library (Stratagene, La Jolla, CA) was screened with a mouse *Cbfa1* DNA fragment (nt 425-884). Phage DNA from two positive clones was digested with *Xba* I and analyzed by Southern blot. We identified a 5.0 kb *Xba* I fragment that was subcloned and sequenced. Sequencing revealed that this fragment contained a 1403 sequence upstream of the previously identified transcription start site [Geoffroy et al., 1998; Ducy et al., 1999; Drissi et al., 2000], the 5′ untranslated region encoded by exon 1.1, the previously described intervening 165 base pairs (bp) mini-intron [Xiao et al., 1998], the remainder of exon 1 (designated exon 1.2) that contains additional 5′ untranslated region and the region coding the N-terminus of the MASNS/p57 isoform (Fig. 1), and a 3.1 kb region corresponding to the first true intron. A 1418 bp sequence containing the 5′-flanking region and 15 bp 3′ to the transcription start site was amplified by polymerase chain reaction (PCR) using a set of primers (5′-GCTCTAGAGCAGCTTGTCGACTAGACATAC-3′) and (5′-GCCTCTAGAGCATGAAGCACTCACACAATC-C-3′) containing *Xba* I sites and the resultant product was subcloned into pluc4 luciferase reporter construct [Ducy et al., 1997] in both direc-tions for subsequent analysis. The promoter analysis programs, Neural Network Promoter Prediction Tool (http://www-hgc.lbl.gov/projects/promoter.html), TFSEARCH (http://pdap1/trc.rwcp.or.jp/research/dp/TRSEARCH-notice.html) and TESS (http://agave.humgen.upenn.edu/utess/tess) were used to identify consensus sequences for potential regulatory elements.

Primer Extension

Primer extension was performed as per the manufacturer's protocol (Promega, Madison,

WI). A γ^{32}P-labeled oligonucleotide primer *Cbfa1*-359.R (5′-GGCTTGTGGTAAGGCCTTC-CTGGCATTCAAG-3′) was annealed to 40 μg of total RNA from MC3T3-E1 for 1 h at 68°C; 50 U of AMV reverse transcriptase (RT) was added, and this was followed by 30 min at 42°C. As a size standard, the cloned 1.4 kb upstream genomic sequence of 5′ end of *Cbfa1* was sequenced with the same oligonucleotide primer.

Reverse Transcription Polymerase Chain Reaction (RT-PCR)

RT-PCR was done using two-step RNA PCR kit (Perkin Elmer Corp., Branchburg, NJ). DNase I-treated total RNA (1.0 μg) was reverse transcribed into cDNA with the reverse primer specified below. The reverse transcription reaction was incubated at 42°C for 60 min. The resulting cDNA was then PCR amplified using various sets of *Cbfa1* primers. Forward primers included *Cbfa1*-453.F (5′-CCAGAGGCT-TAACCTTACAG-3′); Cbfa1-252.F (5′-GACA-GAGGAACACCCATAAG-3′); Cbfa1-123.F (5′-CCACAGTGGTAGGCAGTCCC-3′) Cbfa1+7.F (5′-ATGCTTCATTCGCCTCACAAAC-3′). The reverse primer was *Cbfa1* + 464.R (5′-CCAAAA-GAAGCTTTGCTGAC-3′) located in exon 1. PCR was performed with thermal cycling parameters of 94°C for 30 sec, 60°C for 30 sec, and 72°C for 45 sec for 35 cycles followed by a final extension at 72°C for 7 min. In addition, using a similar protocol [Xiao et al., 1998], the mouse osteocalcin transcript was RT-PCR amplified using the forward primer mOG + 8.F (5′-CAAGTCCCACACAGCAGCTT-3′) and the reverse primer mOG + 378.R (5′-AAAGCC-GAGCTGCCAGAGTT-3′). Mouse β-actin was amplified as a control for the RT-PCR reactions as previously described [Xiao et al., 1998].

Cell Culture

Cos-7, NIH3T3 fibroblast, HEK293, Hepa1-6, and C2C12 cells were maintained in Dulbecco's modified Eagle's medium (DMEM), the C3H10T1/2 pluripotent cell line was maintained in basal medium eagle (BME) and MC3T3-E1 osteoblasts were grown in α-minimum essential medium (MEM) (Gibco-BRL, Gaithersberg, MD). Ros 17/2.8 osteoblasts were grown in one to one mixture of Dulbecco's modified Eagles medium and Nutrient Mixture F-12 (DMEM-F12). All cell lines were supplemented with 10% (vol/vol) FBS and 100 μg/ml penicillin and streptomycin unless otherwise

noted. L929 mouse skin fibroblasts were maintained in MEM/Earle's balanced salt solution (EBSS) (Hyclone, Logan, UT) supplemented with 5% FBS and 100 µg/ml penicillin and streptomycin. All cells were cultured in a humidified incubator with 5% CO_2 at a temperature of 37°C. Recombinant human bone morphogenetic protein-2 (rhBMP-2) was a gift from Genetics Institute, Cambridge, Mass.

DNA Transfection and Reporter Assays

For each transfection, 2×10^5 cells were transfected with the 1.4 kb *Cbfa1* promoter pluc4 plasmid (2.0 µg) and pSV β-gal plasmid (0.5 µg) using Lipofectamine (Gibco BRL, Grand Island, NY). Luciferase activity was measured using the luciferase assay kit (Promega, Madison, WI). β-galactosidase activity was measured by a colorimetric assay (Promega, Madison, WI) and total protein content was determined with a Bio-Rad Protein assay kit (Bio-Rad, Hercules, CA). The luciferase activity was normalized for β-galactosidase activity and total protein amount. The relative luciferase activity was calculated by dividing the normalized luciferase activity by that obtained with the empty pluc4 vector.

Northern Blot and RNase Protection

Total RNA was isolated using TRIzol reagent (Life Technologies, Baltimore, MD). Northern blot analysis was carried out as previously described [Xiao et al., 1998]. Briefly, 15 µg of total RNA was fractionated in a 1.2% formaldehyde agarose gel, then transferred to Nytran Plus membrane. Hybridization was performed with a *Cbfa1* specific probe to exon 1 by modifications of previously described methods [Xiao et al., 1998]. Autoradiography was for three days at −80°C using Biomax film (Kodak, Rochester, NY). RNase protection assays were used to measure *Cbfa1* mRNA expression in various cells. The riboprobe (239,465) consists of a 227 bp fragment of the mouse *Cbfa1* corresponding to 103 bp portion of the mini-intron and 124 bp of exon 1 and the riboprobe (342,546) corresponds to 206 bp of exon 1 (GenBank accession no. AF053948, AF134836, AB013129, AF155360, and AF325502) subcloned into pBS SK (Stratagene, La Jolla, CA). A β-actin riboprobe was transcribed from pTRI-Actin-Mouse (Ambion Inc, Austin, TX) for use as a control for RNA quantity and integr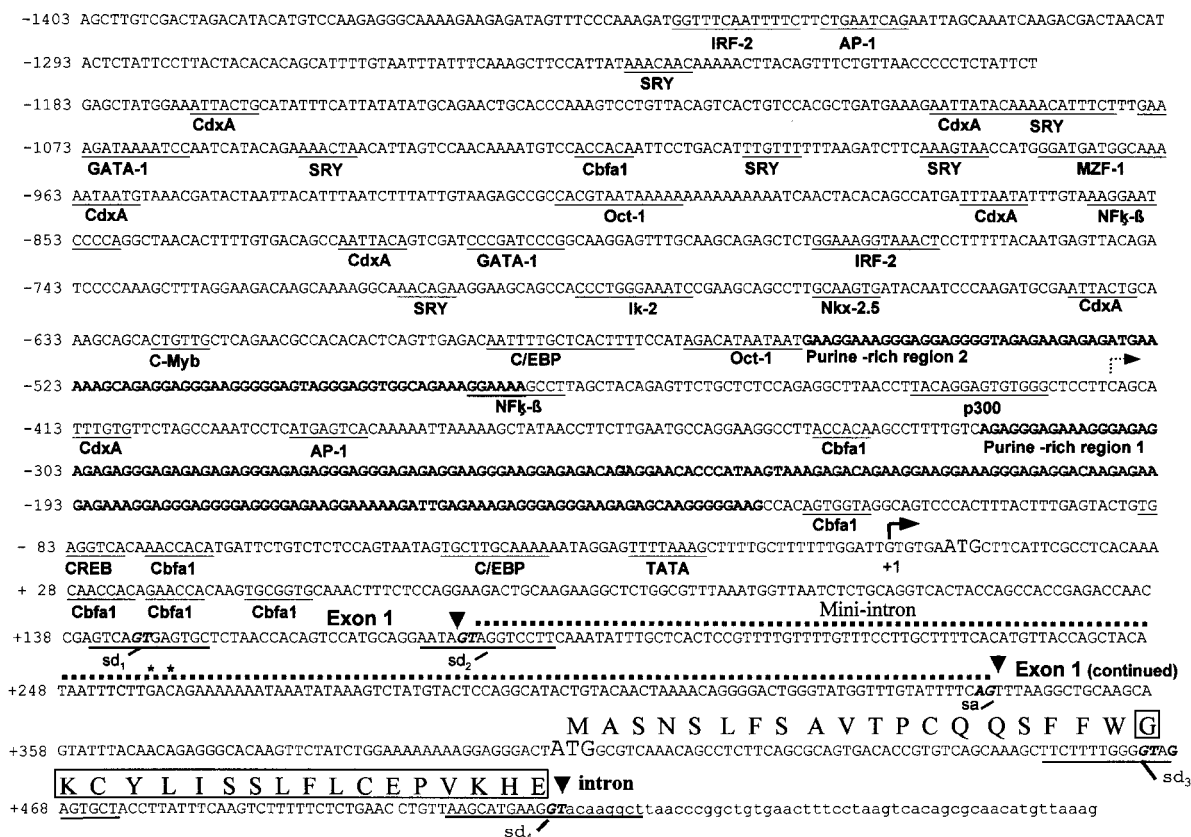ity. Single-stranded anti-sense radiolabeled RNA probes were transcribed from NotI linearized pBSK-Cbfa1 using T7 RNA polymerase and [^{32}P-]UTP (NEN Life Science Products, Boston, MA). All probes were purified on a 5% polyacrylamide gel (250 V for 2 h). Following brief exposure of gel to film, radiolabeled riboprobes were excised and eluted overnight into 350 µl Probe Elution Buffer (Ambion Inc, Austin, TX) at room temperature. RNase protection assays were conducted with a few modifications using the RPA III kit (Ambion Inc, Austin, TX). Total RNA (20 µg) (DNase I-treated) from each cell line and 1,00,000 cpm riboprobe were precipitated and dissolved in 10 µl of Hybridization III Buffer and incubated overnight at 42°C. Products then were digested with RNase A/RNase T1 mix (1:100 dilution) for 30 min, and protected RNA fragments were separated on a 5% denaturing polyacrylamide gel. Radiolabeled RNA Century size markers (Ambion Inc, Austin, TX) and 50 bp DNA ladder (Gibco BRL, Grand Island, NY) were run on the same gel to determine size of products. The resulting gel was transferred to chromatography paper, then wrapped in plastic wrap and exposed to BioMax MS film overnight at −70°C with intensifying screen.

RESULTS

Isolation and Cloning of the Mouse Cbfa1 TATA-Less Promoter

Following determination of the distal exons of the *Cbfa1* gene [Xiao et al., 1998] (Fig. 1), we subcloned and sequenced the 5′-flanking region of the first exon to identify the *Cbfa1* promoter (Fig. 2). Our sequence (Fig. 2; GenBank AF134836) is highly homologous to other GenBank entries, including 92% identity to the human CBFA1 promoter submitted by Ogawa, S. [AB013356]; 98% identity to the overlapping region of the recently published rat upstream promoter [Drissi et al., 2000; AF325502] and 99% identity to the previously published mouse upstream promoters [Fujiwara et al., 1999; AB013129] and the sequence submitted by Chi X.Z. and Bae S.C.[AF155360S1].

Using RNA derived from MC3T3-E1 osteoblasts that are known to express *Cbfa1* [Xiao et al., 1998], the location of the start site within this 5′-flanking region was initially approximated by RT-PCR using primers flanking the initiation sites (Fig. 3A). Forward primers *Cbfa1*+7.F (lane 1), *Cbfa1*-123.F (lane 2) and

```
-1403 AGCTTGTCGACTAGACATACATGTCCAAGAGGGCAAAAGAAGAGATAGTTTCCCAAAGATGGTTTCAATTTTCTTCTGAATCAGAATTAGCAAATCAAGACGACTAACAT
                                                                    IRF-2              AP-1
-1293 ACTCTATTCCTTACTACACACAGCATTTTGTAATTTATTTCAAAGCTTCCATTATAAACAACAAAAACTTACAGTTTCTGTTAACCCCCTCTATTCT
                                                          SRY
-1183 GAGCTATGGAAATTACTGCATATTTCATTATATATGCAGAACTGCACCCAAAGTCCTGTTACAGTCACTGTCCACGCTGATGAAAGAATTATACAAAACATTTCTTTGAA
                CdxA                                                                    CdxA    SRY
-1073 AGATAAAATCCAATCATACAGAAAACTAACATTAGTCCAACAAAATGTCCACCACAATTCCTGACATTTGTTTTTTAAGATCTTCAAAGTAACCATGGGATGATGGCAAA
      GATA-1          SRY                        Cbfa1            SRY              SRY              MZF-1
-963 AATAATGTAAACGATACTAATTACATTTAATCTTTATTGTAAGAGCCGCCACGTAATAAAAAAAAAAAAAAAATCAACTACACAGCCATGATTTAATATTTGTAAAGGAAT
      CdxA                                          Oct-1                             CdxA       NFk-ß
-853 CCCCAGGCTAACACTTTTGTGACAGCCAATTACAGTCGATCCCGATCCCGGCAAGGAGTTTGCAAGCAGAGCTCTGGAAAGGTAAACTCCTTTTTACAATGAGTTACAGA
                  CdxA        GATA-1                                            IRF-2
-743 TCCCCAAAGCTTTAGGAAGACAAGCAAAAGGCAAACAGAAGGAAGCAGCCACCCTGGGAAATCCGAAGCAGCCTTGCAAGTGATACAATCCCAAGATGCGAATTACTGCA
                              SRY              Ik-2              Nkx-2.5                      CdxA
-633 AAGCAGCCACTGTTGCTCAGAACGCCACACACTCAGTTGAGACAATTTTGCTCACTTTTCCATAGACATAATAATGAAGGAAAAGGGAGGAGGGGGTAGAGAAGAGAGATGAA
              C-Myb                                    C/EBP              Oct-1    Purine -rich region 2      ⫶►
-523 AAAGCAGAGGAGGGAAGGGGGAGTAGGGAGGTGGCAGAAAGGAAAAGCCTTAGCTACAGAGTTCTGCTCTCCAGAGGCTTAACCTTACAGGAGTGTGGGCTCCTTCAGCA
                              NFk-ß                                                              p300
-413 TTTGTGTTCTAGCCAAATCCTCATGAGTCACAAAAATTAAAAAGCTATAACCTTCTTGAATGCCAGGAAGGCCTTACCCACAAGCCTTTTGTCAGAGGGAGAAAGGGAGAG
      CdxA        AP-1                                                        Cbfa1         Purine -rich region 1
-303 AGAGAGGGAGAGAGAGAGGGAGAGAGGGAGGGAGAGAGGAAGGGAAGGAGAGACAGAGGAACACCCATAAGTAAAGAGACAGAAGGAAGGAAAGGGAGAGGACAAGAGAA
-193 GAGAAAGGAGGGAGGGGAGGGGAGAAGGAAAAAGATTGAGAAAGAGGGAGGGGAAGAGAGCAAGGGGGAAGCCACAGTGGTAGGCAGTCCCACTTTACTTTGAGTACTGTG
                                                                              Cbfa1    ►
- 83 AGGTCACAAACCACATGATTCTGTCTCTCCAGTAATAGTGCTTGCAAAAAATAGGAGTTTTAAAGCTTTTGCTTTTTTGGATTGTGTGAATGCTTCATTCGCCTCACAAA
      CREB     Cbfa1                            C/EBP        TATA           +1
+ 28 CAACCACAGAACCACAAGTGCGGTGCAAACTTTCTCCAGGAAGACTGCAAGAAGGCTCTGGCGTTTAAATGGTTAATCTCTGCAGGTCACTACCAGCCACCGAGACCAAC
      Cbfa1    Cbfa1    Cbfa1    Exon 1      ▼ ·····················   Mini-intron ···········
+138 CGAGTCAGTGAGTGCTCTAACCACAGTCCATGCAGGAATAGTAGGTCCTTCAAATATTTGCTCACTCCGTTTTGTTTTGTTTCCTTGCTTTTCACATGTTACCAGCTACA
      sd₁   * *                              sd₂
 ···················································································· ▼ Exon 1 (continued)
+248 TAATTTCTTGACAGAAAAAAAATAAATATAAAGTCTATGTACTCCAGGCATACTGTACAACTAAAACAGGGGACTGGGTATGGTTTGTATTTTCAGTTTAAGGCTGCAAGCA
                                                                                          sa
                 M  A  S  N  S  L  F  S  A  V  T  P  C  Q   Q  S  F  F  W  G
+358 GTATTTACAACAGAGGGCACAAGTTCTATCTGGAAAAAAAAGGAGGGGACTATGGCGTCAAACAGCCTCTTCAGCGCAGTGACACCGTGTCAGCAAAGCTTCTTTTGGGGTAG
                                                                                                        sd₃
      K  C  Y  L  I  S  S  L  F  L  C  E  P  V  K  H  E   ▼ intron
+468 AGTGCTACCTTATTTCAAGTCTTTTTCTCTGAAC CTGTTAAGCATGAAGGTacaaggcttaacccggctgtgaactttcctaagtcacagcgcaacatgttaaag
                                       sd₄
```

Fig. 2. Nucleotide sequence of the 5′ flanking region of the *Cbfa1* gene. The sequence encompasses 1868 bp of *Cbfa1*. Exon/intron boundaries are shown by "▼". The first potential transcription start site based on primer extension analysis is shown by the broken arrow "⫶►" and the previously published transcription start site by the solid arrow "►", [Geoffroy et al., 1998]. Consensus sites for known transcription factors are underlined. The sequence is numbered relative to the first potential transcription start site. "sd" and "sa" refer to the splice donor and splice acceptor sites. The broken line (------) depicts the "mini-intron" embedded in exon 1. The novel 18 amino acid insert that is derived by alternative splice site selection is depicted by the box (□). The promoter/reporter construct, p1.4*Cbfa1*-luc, consists of sequence −1403 to +14 (with the exception that the "ATG" in exon 1 is mutated to a "GTG") subcloned into the pLuc4 vector.

Cbfa1-252.F (lane 3) in combination with the reverse primer *Cbfa1* + 464.R in exon 1 generated the predicted size alternatively spliced products (left panel, Fig. 3A) that we previously described in osteoblasts [Xiao et al., 1998]. For controls, we demonstrated using genomic DNA as the template that the respective primers generated their predicted size products (Fig. 3A, right panel). In contrast, the upstream primer *Cbfa1*-453.F (lane 4, left panel, Fig. 3A in combination with *Cbfa1* + 464.R did not amplify the 918 bp product observed with genomic DNA as the template (Fig. 3A, lane 4, right panel), indicating that transcription initiation begins in the segment flanked by primers *Cbfa1*-464.F and *Cbfa1*-252.F (Fig. 3A, left panel), which is upstream of the previously identified transcription initiation site [Geoffroy et al., 1998]. The predominance of transcripts generated by the *Cbfa1* + 7.F compared to the *Cbfa1*-123.F (Fig. 3A, compare lane 1 and 2, left panel), however, suggest that the downstream mRNA cap site located in previous studies is the predominant start site [Geoffroy et al., 1998].

To confirm the presence and more precisely identify the location of our putative upstream transcription start site, primer extension analysis was performed using a synthetic oligonucleotide complementary to the region at +60 to +90 bp (Fig. 3B). Two products were transcribed, the termination points of which were located at 825 and 824 bps upstream from the ATG codon (Fig. 3B), or ~418 bp upstream of the mRNA cap site identified by other investigators [Geoffroy et al., 1998; Drissi et al., 2000]. These data add further support to the presence
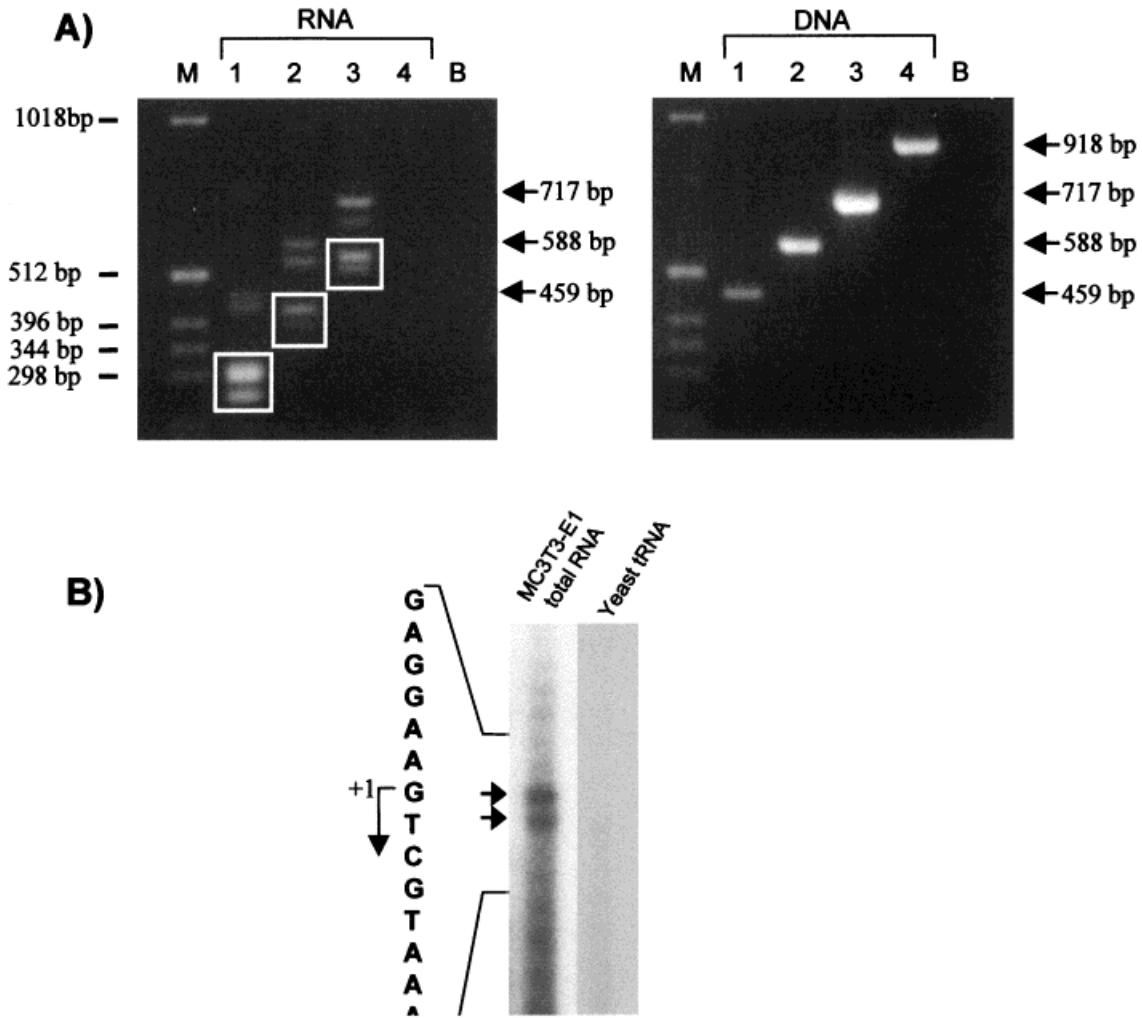
Fig. 3. Determination of the transcription start site of the mouse *Cbfa1* gene. **A:** RT-PCR mapping was performed with 1 μg of total RNA derived from MC3T3-E1 osteoblasts (left panel). PCR using a DNA fragment containing the 5′ end of the *Cbfa1* gene as the template was performed as a control (right panel). The reverse primer, *Cbfa1+464.R* in exon 1 (see Fig. 1), was used in combination with respective forward primers *Cbfa1+7.F* (**lane 1**), *Cbfa1 −123.F* (**lane 2**), *Cbfa1 -252.F* (**lane 3**) and *Cbfa1 −453.F* (**lane 4**). M, 1 kb DNA ladder; B, water blank; bands demarcated by a box "☐" represent *Cbfa1d1* and *Cbfa1d2* alternatively spliced processed mRNA (left panel,

see Fig. 4 for further explanation). The arrows depict the unspliced mRNA containing the mini-intron *Cbfa1u* (left panel) or the anticipated genomic sequence (right panel). A band of ∼426 bp was also identified that may represent retention of the mini-intron and removal of the segment between sd1 and sd2. **B:** Primer extension was performed with 40 μg of total MC3T3-E1 RNA hybridized to γ^{32}P-labeled primer complementary to nucleotide +60 to +90. The position of the 5′-ends was determined by electrophoresis of the primer extension products next to a sequence ladder derived from the same primer.

of transcripts (Fig. 3B) with 5′UTRs that extends beyond the originally defined cap site (Fig. 2).

In addition, in our initial RT-PCR analysis to approximate the start site, we observed additional RT-dependent products in MC3T3-E1 that were larger than the predicted fully processed mRNA (Fig. 3A, lanes 1, 2, and 3). These products, which were identical in size to the genomic DNA derived bands, appeared to correspond with mRNA containing the 165 bp

mini-intron embedded in exon 1 (compare uppermost band in left panel with bands in right panel, Fig. 3A). We also observed RT-dependent bands that are ∼30 bp smaller than the 459 bp product consistent with a mRNA for a previously described alternatively spliced product that utilized the splice donor site, sd2 (Figs. 1 and 2) [Xiao et al., 1998] and two lower bands representing alternatively spliced products *Cbfa1d1* and *Cbfa1d2* in which the mini-intron has been removed (see Fig. 6 below).

We identified a putative TATA-box at position −26 bp upstream of the start site (5′TTTTAAAG) and two purine rich regions at positions −321 to −124 and −559 to −478 separating highly conserved regulatory regions between mouse, human and rat as previously described [Fujiwara et al., 1999; Drissi et al., 2000] (Fig. 2). We also confirmed the presence of multiple conserved recognition motifs identified in these related promoters by other investigators [Fujiwara et al., 1999; Drissi et al., 2000]. Using the criteria of 85% identity to consensus *cis*-acting elements, we found motifs including CREB at position −85/−78, AP-1 at position −391/−383, CBFA at positions −75/−69, −119/−114, −338/333, and −1023/−1018, a NFκβ sites at position −483/−478 and −860/−854, c-Myb site at −626/−C/EPB sites at position −45/−35 and −591/−577 as well as others shown in Figure 2. Notably, the consensus Smad binding sequence CAGACA [Luigi et al., 1998], which mediate bone morphogenic protein effects, was not identified in our sequence. Several homeobox-like binding motifs (designated Cdxa) [Subramanian et al., 1995] and putative SRY *cis*-elements [Pontiggia et al., 1994] were also observed. We also identified potential binding motifs for antagonistic interferon regulatory factors 2 (IRF-2) at −1343/−1332 and −778/−765, which is induced in mineralized bone nodules [Lynch et al., 1998]. A possible DNA binding site for MZF-1 [Morris et al., 1994] was found at positions −1077/−1064.

Cell Specificity and Abundance of Cbfa1 mRNA

To clarify the cell-type specificity and differentiation stage dependency of *Cbfa1* expression, as well as to identify cells for subsequent functional analysis of the 5′-flanking region of the *Cbfa1* gene, we performed Northern and RNase protection analysis of total RNA derived from several different cell lines using a *Cbfa1* specific riboprobes (Fig. 4). For these studies we analyzed mouse fibroblastic cell lines, L929 and NIH3T3, pluripotent C3H10T1/2 cells, and pre-osteoblastic MC3T3-E1 cells. Using a cDNA probe specific for exon 1, we detected a single broad band of the predicted 5.4 kb size, as previously reported [Xiao et al., 1998]. We detected high levels of *Cbfa1* mRNA expression in MC3T3-E1 osteoblasts [Quarles et al., 1997]. In addition, we observed low but detectable

levels of *Cbfa1* transcripts in C3H10T1/2 that are capable of differentiating into osteoblasts under selected conditions, but which do not express osteoblastic features under the conditions studied. We also found evidence for a low abundant transcript in L929 skin fibroblasts, which is of uncertain significance, though dermal cells may possess osteoblastic maturational potential under special circumstances [Yeh et al., 1998]. In contrast, we failed to detect *Cbfa1* transcripts in NIH3T3 fibroblasts by Northern analysis (Fig. 4A).

To confirm the cell-type expression of *Cbfa1*, determine if the mini-intronic genomic sequence is retained in mRNAs and evaluate whether alternatively spliced transcripts are expressed, we performed RNase protection with the riboprobes (239,478) and (342,546). The riboprobe (239,478) detects the mini-intron and regions of exon 1, whereas the riboprobe (342, 546) detects the proximal region of exon 1 containing putative splice donor sites sd3 and sd4 (Fig. 1A). RNase protection confirmed the presence of *Cbfa1* mRNAs in MC3T3-E1, C3H10T1/2 cells, Ros17/2.8 osteoblasts, and L929 fibroblasts and the absence of *Cbfa1* transcripts in Cos-7 cells, HEK293, Hepa1-6, and NIH3T3 fibroblasts (Fig. 4B and C). In addition, we examined C2C12 cells before and after BMP-2 stimulation (Fig. 4B). Unstimulated C2C12 cells had very low levels of expression that was increased after stimulation with BMP-2 (Fig. 4B). Ros17/2.8 osteoblasts expressed *Cbfa1* at levels similar to MC3T3-E1 osteoblasts and BMP-2-stimulated C2C12 cells.

In addition, RNase protect revealed that the single band identified by Northern blot analysis actually is composed of several *Cbfa1* mRNA species, differing in size by up to 81 bps. In this regard, using the riboprobe (239,465) we identified two bands of 227 and 124 bp, consistent with a transcript containing the "mini-intron" in the 5′UTR (designated *Cbfa1u*) and a mRNA species in which this "mini-intronic" sequence has been splice out of the immature mRNA transcript (Fig. 4b). These findings were confirmed by subsequent RT-PCR (see below). Using the riboprobe (342,546), we also identified bands of ∼178 and 124 bp, respectively, corresponding to the mRNAs that appear to be derived from alternative splice site selection (Fig. 4C). Indeed, examination of the sequence revealed a donor splice sites at positions 464 and 517 that accounts for these products (Fig. 2). Translation
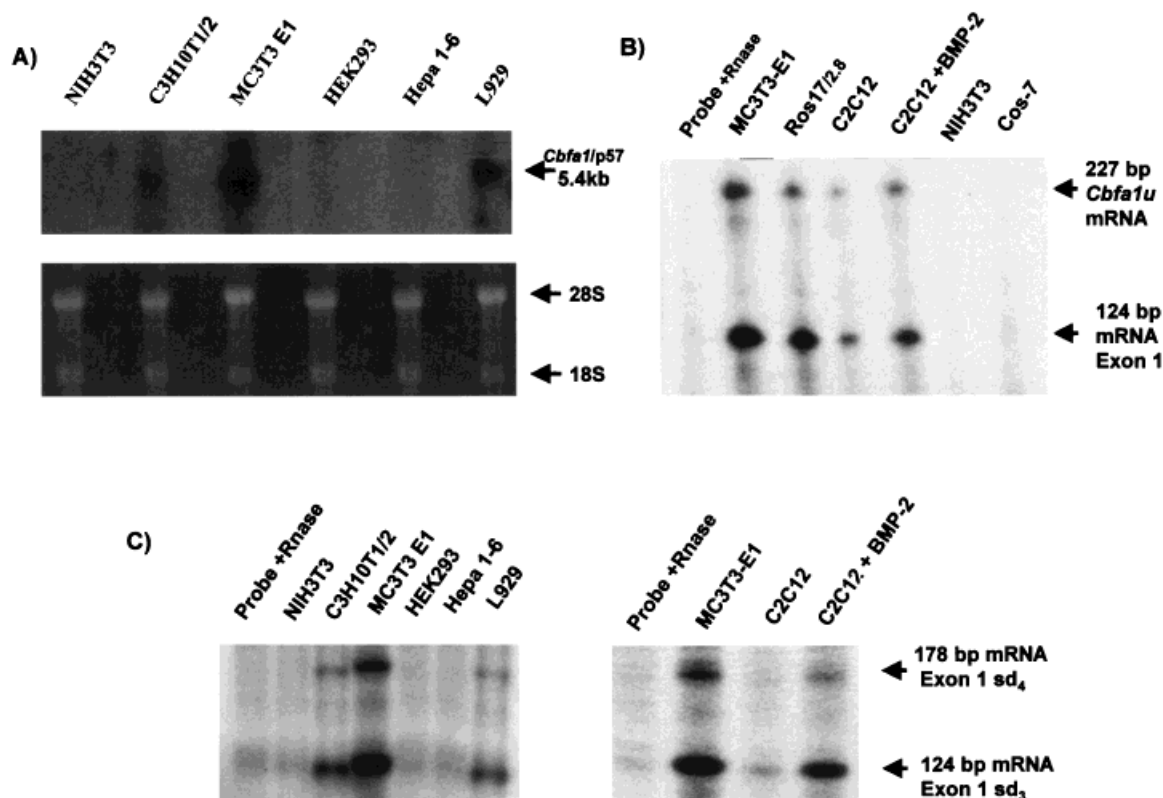
Fig. 4. Northern blot and RNase protection analysis of *Cbfa1* mRNA expression in various cell lines. **A:** Total RNA (15 µg), isolated from NIH3T3 fibroblasts, pluripotent C3H10T1/2 cells, MC3T3-E1 pre-osteoblasts, human embryonic kidney cell line HEK293, hepatocyte derived Hepa1-6 cells and L929 skin fibroblasts was hybridized with a radiolabeled *Cbfa1* probe that recognizes the sequence encoded by exons 1 and 2 (upper panel) or stained with ethidium bromide (lower panel) as a control for RNA loading. Arrows indicate the 5.4 kb *Cbfa1* transcript and the 28S and 18S ribosomal RNA. **B:** RNase protection with a riboprobe (239,465) which detects the mini-intron and exon 1 of *Cbfa1* (see Fig. 1). The upper band represents mRNA in which the "mini-intron" in exon 1 has been retained (*Cbfa1u*) and the lower band is the fully processed mRNA. **C:** RNase protection with a riboprobe (342,546), which overlaps sd3 and sd4 sites in exon 1, detects two products representing the *Cbfa1* and an novel isoform derived from alternative splice site selection. The larger band represents a *Cbfa1* isoform transcript with a 54 bp insert encoding a novel 18 amino acid sequence.

of the sequence encoded by exon 1 shows a potential open reading frame without stop codons that extended beyond the previously reported *Cbfa1* amino acid sequence. This additional sequence is predicted to encode a spliced variant containing an 18 amino acids insert, GKCYLISSLFLCEPVKHE, added to the MASNSLFSAVTPCQQSFFW *Cbfa1* sequence (Figs. 1 and 2). Determining whether this protein variant exists will require additional analysis with isoform specific antibodies, which is beyond the scope of the current studies.

Partial Cell-Restricted Activity of a 1.4-kb Cbfa1 5′ Flanking Region Containing the P1 Promoter Fragment

To examine the transcription of the *Cbfa1* gene, we evaluated cell type-specific expression of our *Cbfa1* promoter fragment. The 5′ flanking region from position. −985 to +433 was linked to the luciferase gene (p1.4*Cbfa1*-luc) and transfected into various cell lines (Fig. 5). We initially compared the activity of p1.4*Cbfa1*-luc in Ros 17/2.8 osteoblasts that express high levels of *Cbfa1* and Cos-7 cells which do not express *Cbfa1* (Fig. 5A). This promoter construct directed high levels of expression (Fig. 5A) in Ros 17/2.8 osteoblasts but not in Cos-7 cells, which are neither osteoblastic nor mesenchymal in origin.

To confirm the specificity of the *Cbfa1* promoter, we evaluated its activity in several other osteoblastic and non-osteoblastic cell lines that express different levels of *Cbfa1* mRNAs. Although we observed high levels of promoter activity in *Cbfa1* expressing MC3T3-E1 osteoblasts similar to Ros 17/2.8 cells, we also found
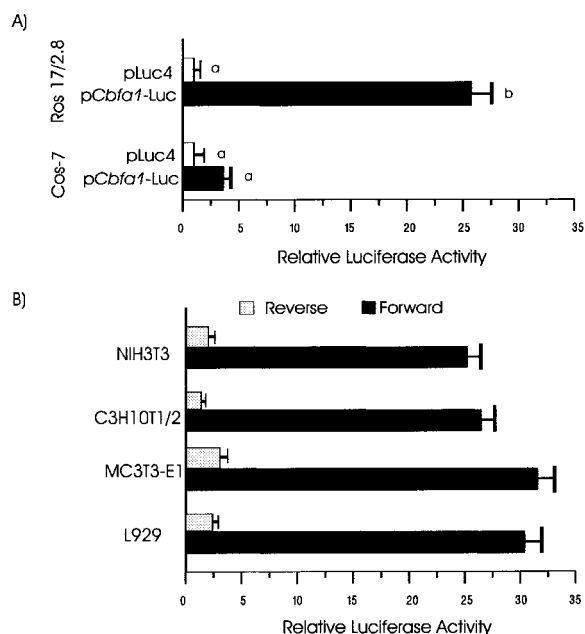
Fig. 5. Functional analysis of the murine *Cbfa1* promoter in different cell lines. We measured the promoter activity of the 5′ flanking region of *Cbfa1* with a transient transfection system, using luciferase as a reporter gene. The constructs used in these studies consisted of fragments of the genomic DNA from the 5′-flanking regions of *Cbfa1* inserted upstream of the firefly luciferase gene in the promoter-less pluc4 vector. **A:** Differential activity in Ros 17/2.8 osteoblasts and Cos-7 cells. A 1.4-kb fragment from the 5′ flanking region of *Cbfa1* containing approximately 1 kb of the promoter in the correct orientation has very strong activity in Ros 17/2.8 osteoblasts (∼25 fold compared to pLuc4 vector alone), but is inactive in Cos-7 cells. Columns sharing the same superscript are not significantly different at $P < 0.05$. **B:** Generalized *Cbfa1* promoter activity in non-osteoblastic cells. The same 1.4-kb fragment from the 5′ flanking region of *Cbfa1* also has very strong activity in NIH3T3 fibroblasts, C$_3$H10T1/2, MC3T3-E1 osteoblasts, and L929 cells. The same construct in the reverse orientation had minimal activity.

high levels of luciferase activity (> 20-fold above controls), driven from the *Cbfa1* promoter in L929 fibroblasts and C3H10T1/2 cells that express low levels of *Cbfa1* mRNA, as well as in NIH3T3 fibroblasts that do not express Cbfa1 transcripts (Fig. 5B). Thus, *Cbfa1* promoter activity of the 1.4 kb construct does not correlate with endogenous expression of processed *Cbfa1* transcripts as assesed by Northern blot analysis (Fig. 4). This indicates that there are likely additional upstream and/or downstream elements that account for the differential transcription of *Cbfa1* in mesenchymal cells, and/or post-transcriptional mechanisms that control cell-restricted expression of *Cbfa1*.

Bone Morphogenetic Protein-2 Does not Stimulate the Cbfa1 Promoter

To explore other cell culture models of osteoblast development that may represent an earlier stage of maturation and that are inducible, we investigated promoter regulation in pluripotent C2C12 cells before and after stimulation with BMP-2 [Katagiri et al., 1994; Tsuji et al., 1998; Lee et al., 2000]. These pluripotent mesenchymal cells become myoblasts in culture unless stimulated toward the osteoblast lineage with factors such as BMP-2 [Katagiri et al., 1994; Nishimura et al., 1998]. We found that the addition of rhBMP-2 to the culture medium of C2C12 cells resulted in a dose-dependent increase in osteocalcin message, consistent with the previously reported ability of such treatment to induce upregulation of the osteoblastic phenotype in these cells [Katagiri et al., 1994] (Fig. 6). In accord with induction of osteoblast differentiation, rhBMP-2 upregulated *Cbfa1 / p57* and its alternatively spliced isoforms (Fig. 6A), similar to the know effect of BMPs to induce Runx2 [Tsuji et al., 1998; Lee et al., 1999, 2000]. In contrast, rhBMP-2 had no effect on transcription of the *Cbfa1* gene as assessed by the activity of our proximal 1.4 kb *Cbfa1* promoter-luciferase construct (Fig. 6B) in C2C12 cells before and after treatment with BMP-2. The levels of *Cbfa1* pre-mRNA (*Cbfa1u*, Fig. 6A) where also not altered by BMP-2 treatment. The failure to detect changes in the promoter activity following BMP-2 treatment indicate the lack BMP-2 responsive elements in the proximal *Cbfa1* promoter. Moreover, induction of osteoblast differentiation by rhBMP-2 appears to stimulate the splicing event that removes the min-intron from exon 1 in the *Cbfa1* gene leading to the increased formation of the isoforms designated as *Cbfa1*d1 and *Cbfa1*d2 alternatively spliced transcripts. The identity of these processed mRNAs (e.g., the respective RT-dependent 294 and 261 bp products, was established by direct sequencing of the subcloned products.

DISCUSSION

The presence of dual promoters and the overall genomic organization of *Cbfa1* is remarkably conserved with other members of this family of transcription factors, including Cbfa1/AML1/Runx1 [Miyoshi et al., 1995; Ghozi et al., 1996; Telfer and Rotherberg, 2001]. The tissue restrict pattern of *Cbfa1*
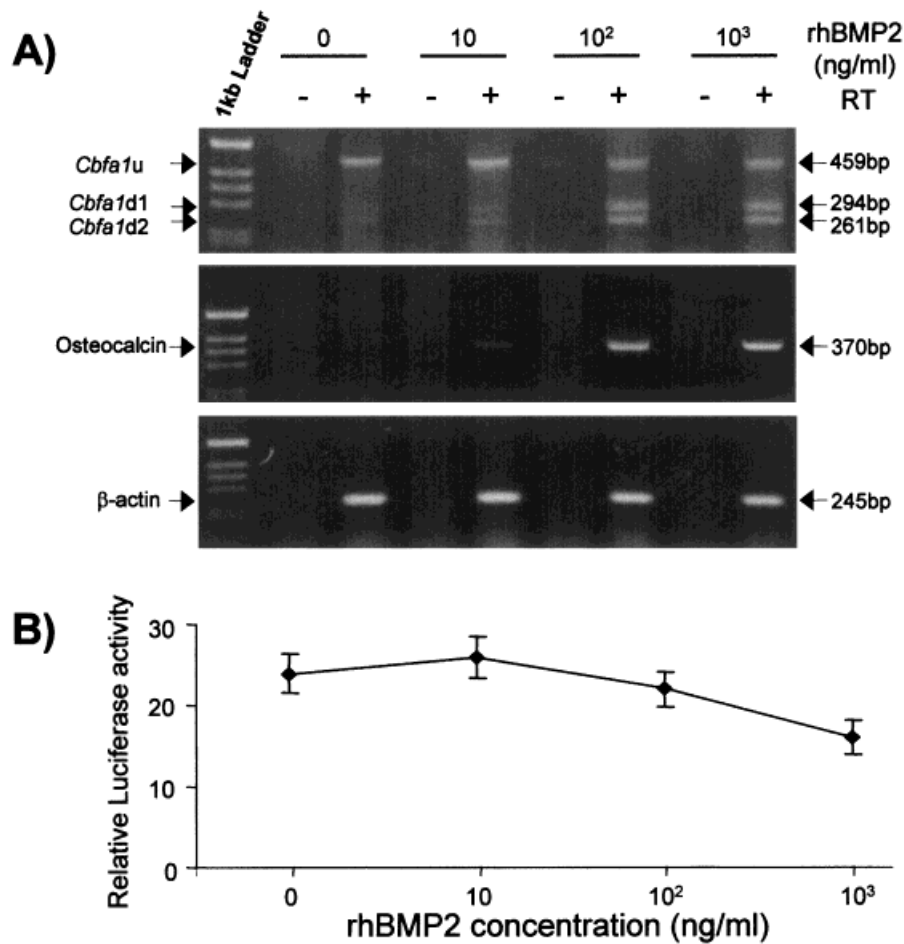
Fig. 6. Dose-dependent effects of rhBMP-2 on *Cbfa1* pre-mRNA splicing and gene transcription in C2C12 cells. **A:** Effect of rhBMP-2 on *Cbfa1* and osteocalcin expression was evaluated by RT-PCR. C2C12 cells were treated with rhBMP2 for 48 hours, and RT-PCR amplification of the transcript corresponding to *Cbfa1u* (459 bp) and *Cbfa1* alternatively spliced products (294 and 261 bp), osteocalcin (370 bp) and β-actin (245 bp) was performed using 1μg total RNA and gene specific primers. −/+ signifies the respective absence and presence of reverse transcriptase (RT). The rhBPM-2 up-regulation of osteocalcin correlated with its stimulation of *Cbfa1* mRNA. **B:** Effect of rhBMP-2 on *Cbfa1* promoter activity was assesed in C2C12 cells transiently transfected with the ~1.0 kb *Cbfa1* promoter-luciferase construct and treated with rhBMP-2 at the indicated concentration for 48 h. rhBMP-2, at concentrations that stimulated expression of mature *Cbfa1* mRNA, failed to stimulate *Cbfa1* promoter activity.

isoform beginning with MASNSL (*Cbfa1*/p57) that regulates bone- and cartilage-related gene expression is believed to be derived from the distal P1 promoter, whereas the expression of the *Cbfa1* isoform beginning with MRIPVD in T-cells may be controlled by the proximal P2 promoter [Drissi et al., 2000]. In the current study, we isolated and characterized a portion of the distal 5′-flanking region of the mouse *Cbfa1* gene and demonstrated that this region contained the P1 promoter (Figs. 1, 2, and 5). Our sequence data are consistent with recent reports characterizing the 5′ regulatory sequences of the mouse *Cbfa1* gene containing

the P1 promoter [Fujiwara et al., 1999; Drissi et al., 2000].

In our studies, however, we found evidence for an additional transcription initiation site in the *Cbfa1* gene (Fig. 3). By RT-PCR and primer extension analysis we identified a putative second transcription start site located 0.4 kb upstream from the previously described transcription start site located near the ATG in exon 1 (Fig. 2) [Geoffroy et al., 1998; Ducy et al., 1999; Fujiwara et al., 1999]. The reason for this discrepancy is not clear but may be related to differences in methodologies. The upstream site may represent a minor alternative start site.

Although this alternative start site would lead to a larger 5′ UTR, which in turn might function to regulate translation [Levanon et al., 2001], confirmation that this alternative site is utilized and its functional significance remain to be established.

Our study also identified a novel N-terminal splice variant involving previously unrecognized splice donor sites (Fig. 2). RNase protection analysis found that a region of exon 1 can give rise to an additional *Cbfa1* isoform through splice site selection containing an 18 amino acids insert, GKCYLISSLFLCEPVKHE, added to the MASNSLFSAVTPCQQSFFW *Cbfa1* sequence (Fig. 1B). The significance of these additional amino acids remains to be clarified. To date, several types of alternative splicing events have been described for *Cbfa1*, including alternative splice site selection and exon skipping [Thirunavukkarasu et al., 1998; Xiao et al., 1998]. We also confirmed the presence of some of these previously described *Cbfa1* isoforms derived from alternative splice site selection (e.g., separate 5′ splice sites are joined to a common 3′ site to generate *Cbfa1d1* and *Cbfa1d2*). Both RT-PCR and RNase protection analysis identified these alternatively spliced products that differ in their 5′UTR by a 33 bp nucleotide insert (Fig. 6). Analysis of our genomic clone also confirmed that *Cbfa1*, similar to the related *Cbfa2* gene, contains a "mini-intron" in the 5′ UTR [Ghozi et al., 1996; Stewart et al., 1997; Kanno et al., 1998; Thirunavukkarasu et al., 1998; Xiao et al., 1998]. Indeed, we found evidence for mRNAs containing this mini-intron in many cell lines (Figs. 4 and 6). Since the strict definition of an intron is that it is not expressed, our RNase protection and RT-PCR detection of the mini-intron indicate that this previously described genomic region is more appropriately designated as part the first exon.

Our results provide insights into the complexity of *Cbfa1/Runx2* gene expression. During embryogenesis *Cbfa1* is expressed at high levels and in a generalized fashion in mesenchymal condensations that give rise to skeletal elements [Ducy et al., 1997; Komori et al., 1997; Mundlos et al., 1997; Otto et al., 1997; Ducy et al., 1999]. Similarly, the luciferase reporter construct that contained ~1.4 kb of the promoter region of the *Cbfa1* gene imparted high levels of expression in osteoblasts and certain mesenchymal-derived cell lines that express fully processed *Cbfa1* messenger RNA, including Ros 17/2.8 osteoblasts, MC3T3-E1 pre-osteoblasts, pluripotent C3H10T1/2 cells, and BMP-2 stimulated C2C12 cells, as well as L929 skin fibroblasts. Expression in non-osteoblastic cells may represent an imprint resulting from the mesenchymal origin of these cells. In addition, we identified putative binding sites for several transcription factors (Fig. 2). Of particular interest are the multiple AML-1/OSE2/CBFA1 binding sites that have been shown to be involved in *Cbfa1* autoregulation of its own expression [Ducy et al., 1999; Fujiwara et al., 1999; Drissi et al., 2000]. Potential SRY-like elements [Bridgewater et al., 1998], which are important in chondrogenesis [Bridgewater et al., 1998] are also present. This may be important since maturation of chondrocytes is disorganized in *Cbfa1*-deficient mice and *Cbfa1* may also play a role in chondrogenesis [Mundlos et al., 1997; Otto et al., 1997]. We failed to identify BMP2 responsiveness in the P1 *Cbfa1* promoter region. Further studies will be needed to determine the physiologically important elements in the *Cbfa1* promoter.

Regardless, additional findings indicate that our existing P1 promoter region may not be sufficient to account for cell-type specific control of *Cbfa1* expression. In this regard, the overall correlation between the *Cbfa1* promoter/reporter activity in specific cell types and the level of *Cbfa1* mRNA expression in various cells was poor (Figs. 4 and 5). Indeed, our promoter construct displayed activity in a variety of cell types that do not express a functional *Cbfa1* transcript by Northern blot and RNase protection analysis (Fig. 5). Since we have isolated only a limited portion of the 5′ flanking region of the *Cbfa1* gene, we cannot exclude the existence of additional upstream or down stream tissue-specific regulatory elements that may be involved in regulating the cell-specific pattern of expression. In support of this possibility, recent studies of a *Cbfa1* promoter construct containing 423 bps of additional upstream sequence demonstrated twice the activity in ROS 17/2.8 osteoblasts compared to C3H10T1/2 cells [Fujiwara et al., 1999]. Down-stream regulatory elements that contribute to cell-specific expression are also possible. In contrast, our *Cbfa1* promoter construct was not functional in non-osteoblastic Cos-7 cells that do not express *Cbfa1* (Fig. 5A). This latter finding may not necessarily reflect an osseous vs. non-

osseous effect, since Cos-7 cells are SV40T antigen transformed and the large T antigen might alter promoter activity.

Additional findings suggest that post-transcriptional processing may be involved in the cell restricted and BMP-2 stimulated expression of mature *Cbfa1* mRNAs. For example, unprocessed *Cbfa1* mRNAs (e.g., retaining the mini-intron) is present in C2C12 cells prior to BMP-2 treatment (Fig. 6). These cells also have the ability to support high levels of *Cbfa1* promoter activity (Fig. 5), but lack evidence for the functional, fully processed *Cbfa1* transcripts. The *Cbfa1u* transcript containing the mini-intron was not present in the absence of reverse transcriptase and thus is not explained by contamination with genomic DNA. While differences in the sensitivity of the detection methods may partially explain these discrepancies, we cannot exclude the possibility that the *Cbfa1* gene is being transcribed at low levels in some cells, but incompletely processed. Indeed, the mini-intron retention in osteoblasts and its removal in response to BMP-2 treatment (Fig. 6), resembles a regulated on/off splicing event similar to what has been described in *Drosophila* of the P transposable element third intron that regulates tissue specific expression [Siebel and Rio, 1990; Lee et al., 1998]. A similar splicing event has been identified in the related *Cbfa 2* gene [Ghozi et al., 1996; Kanno et al., 1998]. Such post-transcriptional regulation is not limited to splicing, since recent studies show that *Cbfa1* is regulated by MAPKs [Xiao et al., 2000].

In conclusion, *Cbfa1* expression is regulated at the transcriptional level and by multiple splicing mechanisms. Further studies are needed to characterize the relative importance of transcriptional and post-transcriptional regulation. Particular studies need to focus on the possibility that splicing of pre-mRNA to mature mRNA plays a role in cell-specific control of *Cbfa1* expression as well as on additional promoter elements that explain the cell restricted and BMP-2 stimulated expression of *Cbfa1*.

REFERENCES

Bridgewater LC, Lefabvre V, de Crombrugghe B. 1998. Chondrocyte-specific enhancer elements in the col11a2 gene resemble the Co12a1 tissue-specific enhancer. J Biol Chem 273:14998–15006.

Drissi H, Luc Q, Shakoori R, Chuva De Sousa Lopes S, Choi JY, Terry AM, Jones S, Neil JC, Lian JB, Stein JL, van Wijnen AJ, Stein GS. 2000. Transcriptional auto-regulation of the bone related CBFA1/RUNX2 gene. J Cell Physiol 184:341–350.

Ducy P, Zhang R, Geoffroy V, Ridall L, Karsenty G. 1997. *Osf2/Cbfa1*: a transcriptional activator of osteoblast differentiation. Cell 89:747–754.

Ducy P, Starbuck M, Priemel M, Shen J, Pinero G, Geoffroy V, Amling M, Karsenty G. 1999. A *Cbfa1*-dependent genetic pathway controls bone formation beyond embryonic development. Genes & Development 13:1025–1036.

Fujiwara M, Tagashira S, Harada H, Ogawa S, Katsumata T, Nakatsuka M, Komori T, Takada H. 1999. Isolation and characterization of the distal promoter region of mouse Cbfa1. Biochim Biophys Acta 1446:265–272.

Geoffroy V, Corral DA, Zhou L, Lee B, Karsenty G. 1998. Genomic organization, expression of the human CBFA1 gene, and evidence for an alternative splicing event affecting protein function. Mamm Genome 9:54–57.

Ghozi MC, Bernstein Y, Negreanu V, Levanon D, Groner Y. 1996. Expression of the human acute myeloid leukemia gene AML1 is regulated by two promoter regions. Proc Natl Acad Sci USA 93:1935–1940.

Harada H, Tagashira S, Fukiwara M, Osawa S, Katsumata T, Yamaguchi A, Komori T, Nakatsuka M. 1999. *Cbfa1* isoforms exert functional differences in osteoblast differentiation. J Biol Chem 274:6972–6978.

Kanno T, Kanno Y, Chen LF, Ogawa E, Kim WY, Ito Y. 1998. Intrinsic transcriptional activation-inhibition domains of the polyomavirus enhancer binding protein 2/core binding factor α subunit revealed in the presence of the β subunit. Mol Cell Biol 18:2444–2454.

Katagiri T, Yamaguchi A, Komaki M, Abe E, Takahashi N, Ikeda T, Rosen V, Wozney JM, Fujisawa-Sehara A, Suda T. 1994. Bone morphogenetic protein-2 converts the differentiation pathway of C2C12 myoblasts into the osteoblast lineage. J Cell Biol 127:1755–1766.

Komori T, Yagi H, Nomura S, Yamaguchi A, Sasaki A, Deguchi K, Shimizu Y, Bronson RT, Gao YH, Inada M, Sato M, Okamoto R, Kitamura Y, Yoshiki S, Kishimoto T. 1997. Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. Cell 89:755–764.

Lee CC, Beall EL, Rio DC. 1998. DNA binding by the KP repressor protein inhibits P-element transposase activity in vitro. EMBO J 17:4166–4174.

Lee MH, Javed A, Kim HJ, Shin HI, Gutierrez S, Choi JY, Rosen V, Stein JL, van Wijnen AJ, Stein GS, Lian JB, Ryoo HM. 1999. Transient up-regulation of CBFA1 in response to bone morphogenetic protein-2 and transforming growth factor β1 in C2C12 myogenic cells coincides with suppression of the myogenic phenotype but is not sufficient for osteoblast differentiation. J Cell Biochem 73:114–125.

Lee KS, Kim HJ, Li QL, Chi XZ, Ueta C, Komori T, Wozney JM, Kim EG, Choi JY, Ryoo HM, Bae SC. 2000. Runx2 is a common target of transforming growth factor β1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific gene expression in the pluripotent mesenchymal precursor cell line C2C12. Mol Cell Biol 20:8783–8792.

Levanon D, Glusman G, Bangsow T, Ben-Asher E, Male DA, Avidan N, Bangsow C, Hattori M, Taylor TD, Taudien S, Biechschmidt K, Shimizu N, Rosenthal A, Sakaki Y, Lancet D, Groner Y. 2001. Architecture and

anatomy of the genomic locus encoding the human leukemia-associated transcription factor RNX1/AML1. Gene 262:23–33.

Luigi J, Jonk C, Itoh S, Heldin C-H, Dijke P, Kruijer W. 1998. Identification and functional characterization of a Smad binding element (SBE) in the JunB promoter that acts as a transforming growth factor-, activin, and bone morphogenetic protein-inducible enhancer. J Biol Chem 273:21145–21152.

Lynch MP, Capparelli C, Stein JL, Stein GS, Lian JB. 1998. Apoptosis during bone-like tissue development in vitro. J Cell Biochem 68:31–49.

Morris JF, Hromas R, Rauscher FJ. 1994. Characterization of the DNA-binding properties of the myeloid zinc finger protein MZF1: two independent DNA-binding domains recognize two DNA consensus sequences with a common G-rich core. Mol Cell Biol 14:1786–1795.

Mundlos S, Otto F, Mundlos C, Mulliken JB, Aylsworth AS, Albright S, Lindhout D, Cole WG, Henn W, Knoll JH, Owen M J, Mertelsmann R, Zabel BU, Olsen BR. 1997. Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. Cell 89:773–779.

Miyoshi H, Ohira M, Shimizu K, Matani K, Hirai H, Imai T, Yokoyama K, Soeda E, Ohki M. 1995. Alternative splicing and genomic structure of the AML1 gene involved in acute myeloid leukemia. Nucleic Acids Res 23:2762–2769.

Nishimura R, Kato Y, Chen D, Harris SE, Mundy GR, Yoneda T. 1998. Smad5 and DPC4 are key molecules in mediating BMP-2-induced osteoblastic differentiation of the pluripotent mesenchymal precursor cell line C2C12. J Biol Chem 273:1872–1879.

Ogawa E, Maruyama M, Kagoshima H, Inuzuka M, Lu J, Satake M, Shigesada IY. 1993. PEBP2/PEA2 represents a family of transcription factors homologous to the products of the Drosophila runt gene and the human AML1 gene. Proc Natl Acad Sci USA 90:6859–6863.

Otto F, Thornell AP, Crompton T, Denzel A, Gilmour KC, Rosewell IR, Stamp GW, Beddington RS, Mundlos S, Olsen BR, Selby PB, Owen MJ. 1997. *Cbfa1*, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. Cell 89:765–771.

Pontiggia A, Rimini R, Harley VR, Goodfellow P N, Lovell-Badge R, Bianchi ME. 1994. Sex-reversing mutations affect the architecture of SRY-DNA complexes. EMBO J 13:6115–6124.

Quarles LD, Siddhanti SR, Medda S. 1997. Developmental regulation of osteocalcin expression in MC3T3-E1 osteoblasts: minimal role of the proximal E-box *cis*-acting promoter elements. J Cell Biochem 65:11–24.

Satake M, Nomura S, Yamaguchi-Iwai Y, Takahama Y, Hashimoto Y, Niki M, Kitamura Y, Ito Y. 1995. Expression of the runt domain-encoding PEBP2α genes in T cells during thymic development. Mol Cell Biol 15:1662– 1670.

Siebel CW, Rio DC. 1990. Regulated splicing of the *Drosophila P* transposable element third intron in vitro: somatic repression. Science 248:1200–1208.

Stewart M, Terry A, Hu M, O'Hara M, Blyth K, Baxter E, Cameron E, Onions DE, Neil J C. 1997. Proviral insertions induce the expression of bone-specific isoforms of PEBP2αA (CBFA1): evidence for a new myc collaborating oncogene. Proc Natl Acad Sci USA 94:8646–8651.

Subramanian V, Meyer BI, Gruss P. 1995. Disruption of the murine homeobox gene Cdx1 affects axial skeletal identities by altering themesodermal expression domains of Hox genes. Cell 83:641–653.

Telfer JC, Rotherberg EV. 2001. Expression and function of a stem cell promoter for the murine CBFα2 gene: distinct roles and regulation in natural killer and T cell development. Dev Biol 229:363–382.

Thirunavukkarasu K, Mahajan M, McLarren KW, Stifani S, Karsenty G. 1998. Two domains unique to osteoblast-specific transcription factor *Osf2 / Cbfga1* contribute to its transactivation function and its inability to heterodimerize with Cbfab. Mol Cell Biol 18:4197–4208.

Tsuji K, Ito Y, Noda M. 1998. Expression of the PEBP2αA/ AML3/CBFA1 gene is regulated by BMP4/7 heterodimer and its overexpression suppresses type I collagen and osteocalcin gene expression in osteoblastic and nonosteoblastic mesenchymal cells. Bone 2:87–92.

Xiao ZS, Thomas R, Hinson TK, Quarles LD. 1998. Genomic structure and isoform expression of the mouse, rat and human *Cbfa1 / Osf2* transcription factor. Gene 214:187– 197.

Xiao G, Jiang D, Thomas P, Benson MD, Guan K, Karsenty G, Franceschi RT. 2000. MAPK pathways activate and phosphorylate the osteoblast-specific transcription factor, Cbfa1. J Biol Chem 275:4453–4459.

Yeh G, Olmsted E, Mathur S, Shore E, Gannon F, Ulied A, Audi L, Kaplan F. 1998. Mis-expression of the Cbfa-1 gene in dermal fibroblasts from a patient with an atypical form of progressive osseous heteroplasia. J Bone Miner Res 23:S648.