TRANSGENIC MICE*  

Molecular Cloning, Characterization, and Promoter Analysis of the Mouse Crp2/SmLim Gene

PREFERENTIAL EXPRESSION OF ITS PROMOTER IN THE VASCULAR SMOOTH MUSCLE CELLS OF TRANSGENIC MICE*

(Received for publication, January 16, 1998)

Shaw-Fang Yet, Sara C. Folta, Mukesh K. Jain, Chung-Ming Hsieh, Koji Maemura, Matthew D. Layne, Dorothy Zhang, Pooja B. Marria, Masao Yoshizumi, Michael T. Chin, Mark A. Perrella, and Mu-En Lee

From the Cardiovascular Biology Laboratory, Harvard School of Public Health, Boston, the Department of Medicine, Harvard Medical School, Boston, and the Cardiovascular and Pulmonary Divisions, Brigham and Women’s Hospital, Boston, Massachusetts 02115

Several members of the LIM protein family play important roles in development and differentiation. We recently isolated a rat cDNA encoding a new member of this family, CRP2/SmLim, that contains two LIM domains and is expressed preferentially in vascular smooth muscle cells (VSMC). To study the molecular mechanisms that regulate VSMC-specific transcription of the Crp2/SmLim gene, we cloned the cDNA and gene of mouse Crp2/SmLim. Mouse Crp2/SmLim is a single copy gene of six exons and five introns spanning approximately 20 kilobases of genomic DNA. By 5′-rapid amplification of cDNA ends and S1 nuclease protection assay, we determined that the transcription start site is an A residue 80 base pairs 5′ of the translation initiation codon. A TATA-like sequence is located 27 base pairs 5′ of the transcription start site, and there are potential cis-acting elements (GATA, Sp1, AP-2, E box, CCAC box, and GArC motif) in the 5′-flanking sequence. In transient transfection assays in rat aortic smooth muscle cells in primary culture, 5 kilobases of the Crp2/SmLim 5′-flanking sequence generated a high level of luciferase reporter gene activity. By deletion analysis and gel mobility shift assay, we found that the region between bases −74 and −39 of this 5 kilobase DNA fragment binds Sp1 and confers basal promoter activity in the Crp2/SmLim gene. In vitro, the 5-kilobase fragment was active in multiple cell types. In vivo, however, the 5-kilobase fragment directed high level expression of the lacZ reporter gene preferentially in the VSMC of transgenic mice, indicating the presence of VSMC-specific element(s) in this fragment.

Members of the LIM protein family play important roles in determining cell lineage and in regulating cell differentiation (1–3). The LIM domain is defined by a 50–60 amino acid consensus sequence (CX2CX17-21HX6CX4(C/D/H)2(4) that serves as a protein interaction domain (5). There are three classes of LIM proteins. Class 1 proteins (LIM-HD) contain two LIM domains and a DNA-binding homeodomain (6–8). Class 2 proteins (LIM-only) contain one or more LIM domains but lack the DNA-binding homeodomain (2). Class 3 proteins (LIM-K) contain LIM domains and a protein kinase domain (9–11). By gene disruption in mice, two members of the LIM-only protein family, RBNT-2 (4) and muscle LIM protein (MLP),1 also known as cysteine-rich protein 3 (CRP3) (12, 13), have been shown to be important in erythroid cell development and striated muscle cell differentiation, respectively. Recently, we cloned a rat LIM-only protein and named it SmLim for its preferential expression in vascular smooth muscle cells (VSMC) (14). Because cDNAs encoding the SmLim homologue CRP2 have also been cloned in chicken and quail (15, 16), we refer to the gene as Crp2/SmLim.

Members of the CRP family of the LIM-only proteins are characterized by two LIM domains that are each followed by a conserved glycine-rich repeat (1, 14, 16–18). The CRP family includes CRP1, CRP2/SmLim, and CRP3/MLP. In adults, CRP1 is expressed in many tissues and cell types. CRP2/SmLim is expressed mainly in VSMC, and CRP3/MLP is expressed only in striated muscle cells (12, 14, 19). CRP3/MLP-deficient mice develop heart failure soon after birth (13), most likely because of a disruption of the cardiomyocyte cytoarchitecture. Although the biological functions of CRP1 and CRP2/SmLim have not been elucidated completely, it has been suggested that they play a similar role in maintaining the cytoarchitecture of smooth muscle cells and therefore affect smooth muscle cell differentiation (18). We have shown that stimulating VSMC in culture with the potent mitogen platelet-derived growth factor-BB markedly down-regulates CRP2/SmLim mRNA (14). In vivo, CRP2/SmLim mRNA levels decrease as VSMC change from a differentiated to a dedifferentiated phenotype in response to vascular injury (14). This down-regulation suggests that CRP2/SmLim may be involved in VSMC growth and differentiation.

As a first step toward elucidating the biological function of CRP2/SmLim and determining the molecular mechanisms regulating its expression, we cloned and characterized the mouse

1 The abbreviations used are: MLP, muscle LIM protein; CRP, cysteine-rich protein; VSMC, vascular smooth muscle cells; kb, kilobases; 5′-RACE, 5′-rapid amplification of cDNA ends; bp, base pairs; RASMC, rat aortic smooth muscle cells; PCR, polymerase chain reaction; MEF2, myocyte-specific enhancer factor 2; Pipes, pipеразин-N,N′-bis/2-ethanesulfonic acid; kb, kilobase.

* This work was supported by a grant from Bristol-Myers Squibb and by National Institutes of Health Grants HL03194 (to M. A. P.) and GM53249 (to M.-E. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EBI Data Bank with accession number(s) AF037208 for the cDNA and AF037209 for the 5′-flanking region.

To whom correspondence should be addressed: Cardiovascular Biology Laboratory, Harvard School of Public Health, 677 Huntington Ave., Boston, MA 02115. Tel.: 617-432-4994; Fax: 617-432-0031; E-mail: lee@cvlab.harvard.edu.
The Crp2/SmLim gene and promoter. Crp2/SmLim is a single copy gene composed of six exons and five introns spanning approximately 20 kilobases (kb). We determined its transcription start site by 5′-rapid amplification of cDNA ends (5′-RACE) and S1 nuclease protection assays. By transient transfection assays, we found that the 5-kb 5′-flanking sequence confers a high level of transcriptional activity in VSMC in culture. Furthermore, this 5-kb 5′-flanking sequence directs high level expression of the lacZ reporter gene mainly in the VSMC of transgenic mice.

**EXPERIMENTAL PROCEDURES**

**Isolation of Mouse CRP2/SmLIM cDNA**—The coding region of the mouse CRP2/SmLIM cDNA was amplified from mouse aorta RNA by the reverse transcription polymerase chain reaction (PCR) as described (20). The forward primer (5′-CCTGTCTGGGCGTGTTGAAGAAATGGGAAAGGACATGGGAAGCTTCTCAATAAA) and the reverse primer (5′-CTACTGAGCATGAACAAGG) were designed according to the rat CRP2/SmLIM cDNA open reading frame sequence (14). The PCR products were subcloned into the pCR1 vector (Invitrogen, San Diego, CA). To isolate the 5′ end of the mouse CRP2/SmLIM cDNA, we performed 5′-RACE (Life Technologies, Inc.) with mouse aorta RNA. The sequence of the 3′-untranslated region was obtained by 3′-RACE (Life Technologies) with mouse aorta RNA.

Nucleotide sequences were determined by the dideoxy chain termination method with Sequenase version 2 (Amersham Pharmacia Biotech) and on an automated DNA Sequencer (LICOR, Lincoln, NE) according to the manufacturer's instructions. dITP and deaza-dGTP sequencing protocols were used to resolve compression artifacts in the highly GC-rich regions. Sequence analysis was performed with the GCG software package (Genetics Computer Group, Madison, WI).

**Southern Blot Analysis**—High molecular weight genomic DNA was prepared from mouse tail biopsies (21). Genomic DNA (10 μg) was digested with restriction enzymes (Lloc, Lincoln, NE) according to the manufacturer's instructions. dITP and deaza-dGTP sequencing protocols were used to resolve compression artifacts in the highly GC-rich regions. Sequence analysis was performed with the GCG software package (Genetics Computer Group, Madison, WI).

**Cloning the Mouse Crp2/SmLim Gene**—The full-length CRP2/SmLim cDNA was labeled by random priming (Stratagene) and used as a probe to screen a mouse 129SvJ l DASH II genomic library (Stratagene, La Jolla, CA) and an automated DNA sequencing machine (Lloc, Lincoln, NE) according to the manufacturer's instructions. dITP and deaza-dGTP sequencing protocols were used to resolve compression artifacts in the highly GC-rich regions. Sequence analysis was performed with the GCG software package (Genetics Computer Group, Madison, WI).
annealed to the alkaline-denatured plasmid extended with the Klenow fragment and then digested with to generate a 410-bp fragment (bp) was labeled with \(^{32}\text{P}\) by random priming and used to screen a mouse 129SVJ genomic library in AFXII (Stratagene). About 1.5 million phage were plated, transferred to nitrocellulose filters, and screened at high stringency. Four clones were isolated that contained the Crp2/SmLim promoter sequence.

**S1 Nuclease Protection Analysis**—The Crp2/SmLim transcription start site was determined by S1 nuclease protection assay. An end-labeled antisense oligonucleotide corresponding to bp +79 to +100 was annealed to the alkaline-denatured plasmid –438/+79, which contains 438 bp of the Crp2/SmLim promoter plus exon 1. The primer was extended with the Klenow fragment and then digested with NdeI to generate a 16-bp fragment (~332 to +79). The purified, labeled probe (5×10⁶ cpm) was then hybridized to 30 μg of each RNA sample in 20 μl of hybridization buffer containing 80% formamide, 40 mM Pipes, 0.4 mM NaCl, and 1 mM EDTA at 30 °C for 16 h. The DNA-RNA hybrid was mixed with 300 μl of nuclease S1 mapping buffer (0.28 M NaCl, 0.05 M sodium acetate (pH 4.5), 4.5 mM MnSO₄, and 20 μg/ml herring sperm DNA) and 100 units of S1 nuclease (Promega, Madison, WI). The mixture was incubated at 30 °C for 60 min. The reaction was terminated with 80 μl of stop buffer containing 4 mM ammonium acetate, 20 mM EDTA, and 40 μg/ml RNA. Final reaction products were analyzed on an 8% sequencing gel.

**Cell Culture**—Rat aortic smooth muscle cells (RASMC) were harvested from male Sprague-Dawley rats (200–250 g) by enzymatic dissociation according to the method of Günther et al. (25). RASMC were cultured in Dulbecco’s modified Eagle’s medium (JRH Biosciences, Lenexa, KS) supplemented with 10% fetal bovine serum (HyClone, Logan, UT), penicillin (100 units/ml), streptomycin (100 μg/ml), and 25 mM HEPES (pH 7.4) (JRH Biosciences). Cells were passaged every 3–5 days, and experiments were performed on cells 5–7 passages from primary culture. HeLa human epithelial carcinoma cells, 3T3 mouse fibroblast cells, and HepG2 human hepatocellular carcinoma cells were grown in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum, penicillin (100 units/ml), and streptomycin (100 μg/ml).

**Plasmid Constructs**—Reporter constructs containing fragments of the Crp2/SmLim 5’-flanking region were inserted into pGL2-Basic and named according to the length of the fragment relative to the transcription start site. A DNA fragment containing 4855 to +79 was amplified from phage DNA by PCR with the Expand Long Template PCR kit (Boehringer Mannheim). The fragment was then ligated into the pGL2-Basic vector to generate the 4855SmLim luciferase reporter construct. PCR was used to generate fragments with increasingly shorter segments of Crp2/SmLim promoter sequence. These constructs shared a common 5’ end at bp +10 but differed at the 5’ end (at bp +438, +91, and +74). Luciferase plasmids were generated by religation after SacI digestion of plasmid p4838SmLim. Luciferase plasmids were generated by cloning annealed, double-stranded complementary oligonucleotides into pGL2-Basic. The 5’-flanking region of the mouse Crp2/SmLim gene spanning bp –4855 to +79 was also subcloned into the promoterless vector pPD46.21 (24) to generate the transgenic construct –4855SmLim-lacZ.

**Transient Transfection—RASMC** were transfected by the DEAE-dextran method (25). Approximately 5×10⁵ cells were plated onto 100-mm tissue culture dishes and allowed to grow for 48–72 h (until 80–90% confluent). To correct for differences in transfection efficiency, the luciferase plasmids and pOPRSVICAT (Stratagene) (5 μg each) were added to RASMC in a solution containing 500 μg/ml DEAE-dextran. After 4 h, cells were washed twice according to the manufacturer’s instructions (Promega) 48 h after transfection. Luciferase activity was measured with an EG&G AutoLumat LB953 luminometer (Gaithersburg, MD) and the Promega luciferase assay system. Chloramphenicol acetyltransferase activity was assayed by a modified two-phase fluor diffusion method (26). The ratio of luciferase activity to chloramphenicol acetyltransferase activity in each sample served as a measure of normalized luciferase activity.

**Gel Mobility Shift Assay**—Probes were made from double-stranded oligonucleotides corresponding to bp –74 to –39 of Crp2/SmLim which contains two Sp1 sites. The probes were radiolabeled as described (27, 28). A typical binding reaction mixture contained DNA probe at 50,000 cpm, 1 μg of poly(dI-dC), 25 mM HEPES (pH 7.9), 40 mM KCl, 3 mM MgCl₂, 1 mM dithiothreitol, 10% glycerol, and 5 μg of RASMC nuclear extract in 25 μl of binding buffer. The reaction mixture was incubated at room temperature for 20 min and analyzed by native polyacrylamide gel electrophoresis in 0.25× TBE (22 mM Tris, 22 mM boric acid, 0.5 mM EDTA). To determine the specificity of the DNA-protein complexes, we performed competition assays with unlabeled double-stranded oligonucleotides encoding the Crp2/SmLim promoter sequence. ATF sequence, and 5’ extension (T-CCCGCGCCTGGGCGCCTGATG-3′). To characterize the specific DNA-binding protein, we incubated nuclear extracts with antibodies to Sp1 and AP2 (Santa Cruz) for 3 h at 4 °C before adding the probe.

**Transgenic Mouse**—The –4855SmLim-lacZ transgenic construct was digested with Psil and NotI to generate a 8-kb fragment containing the 5-kb promoter, lacZ reporter gene, and SV40 polyadenylation sequence. The purified 8-kb fragment was injected into the pronuclei of fertilized FVB mouse eggs (Brigham and Women’s Hospital, Core Transgenic Mouse Facility). Transgenic mice harboring the Crp2/SmLim promoter-lacZ gene were identified by PCR and Southern blot analysis with genomic DNA prepared from tail biopsies of potential founders. Two independent founder lines were identified. Transgenic embryos were fixed and stained for β-galactosidase activity. To make the vasculature more visible before photography, we clarified the stained, whole embryos by soaking them serially in glycerol at concentrations ranging from 30 to 80%. Stained embryos were also embedded in paraffin, sectioned on a microtome at a thickness of 5 μm, and counterstained with nuclear fast red according to standard procedures (29).
Isolation of Mouse CRP2/SmLIM cDNA—By using mouse aorta RNA as template, we isolated the mouse CRP2/SmLIM cDNA by reverse transcription PCR in conjunction with 5'- and 3'-RACE. The mouse CRP2/SmLIM cDNA encodes a 193-amino acid protein that contains two LIM domains separated by a glycine-rich region (Fig. 1A).

Crp2/SmLim Is a Single Copy Gene—Hybridization of mouse genomic DNA digested with SacI, BamHI, HindIII, EcoRI, or PstI to a probe encoding the CRP2/SmLIM open reading frame revealed a simple pattern of bands (data not shown). To exclude the possibility that CRP2/SmLIM cross-hybridized to other CRP protein family genes, we also hybridized a parallel blot with a probe containing a part of the CRP2/SmLIM 3'-untranslated region that has little homology to other CRP cDNA sequences. The 3'-untranslated region probe hybridized to only one band (data not shown). These data indicate that in the mouse genome Crp2/SmLim is encoded by a single copy gene.

Genomic Organization of Mouse Crp2/SmLim Gene—By using the full-length CRP2/SmLIM cDNA as a probe to screen a library prepared from mouse genomic DNA (129SvJ), we isolated several overlapping clones encoding the Crp2/SmLim genomic sequence. The mouse Crp2/SmLim gene is organized...
into six exons and five introns spanning approximately 20 kb of genomic DNA (Fig. 1B). All exon/intron boundaries were determined by DNA sequencing, and intron sizes were determined by PCR with oligonucleotides derived from flanking exons (Fig. 1A). All assigned exon/intron boundaries agree with consensus 5'-GT and 3'-AG splicing sequences (30). The average length of the Crp2/SmLim exons, 145 bp, is consistent with the length of exons in general, 137 bp (31). Exon 1 contains the 5'-untranslated region. Exon 2 contains the entire 5'-finger and the first cysteine of the 3'-finger of the first LIM domain. Exon 3 contains the rest of the 3'-finger. Exon 4 encodes the first two cysteines of the 5'-finger of the second LIM domain, and exon 5 contains the rest of the 5'-finger and most of the 3'-finger except the last cysteine. Exon 6 contains the remainder of the coding region and the 3'-untranslated region. At the 3' end of the Crp2/SmLim gene, exon 6 contains a consensus polyadenylation signal (AATAAA) motif. Poly(A) addition occurs 11 bp downstream of this motif.

Identification of the Crp2/SmLim Transcription Start Site and Potential cis-Acting Sequences—The transcription start site identified by 5'-RACE was confirmed further by using an end-labeled CRP2/SmLIM DNA fragment (nucleotides 2332 to 179) in an S1 nuclease protection assay (Fig. 2A). Protected fragments were observed when the probe was incubated with total RNA prepared from mouse aortic smooth muscle cells in primary culture but not when it was incubated with tRNA.
total RNA from mouse skeletal muscle cells (Fig. 2B). The transcription start site was determined to be a A residue 80 bp 5’ of the translation initiation codon, consistent with the transcription start site was determined to be an A residue 80 bp 5’-RACE. The 5’-CA-3’ nucleotide pair at this position is the most common type of eukaryotic transcription start site (32).

The proximal 5′-flanking region of Crp2/SmLim is characterized by regions rich in G and C (Fig. 3). The Crp2/SmLim transcription start site contains a pyrimidine-rich sequence that resembles the initiator motif (YAYTCCYY) (33). This initiator element, which is often associated with TATA-less promoters, is sufficient to initiate transcription (34). However, we also located a TATA-like sequence (TTTAAA) 27 bp 5′ of the transcription start site. A comparison of the 5′-flanking sequence with sequences in the Transcription Factors Data Base revealed numerous Sp1 sites and six consensus AP-2 sites at positions that resemble the initiator motif (YAYTCCYY) (33). This initiator element, which is often associated with TATA-less promoters, is sufficient to initiate transcription (34). However, we also located a TATA-like sequence (TTTAAA) 27 bp 5′ of the transcription start site.

Analysis of the Promoter Activity of the 5′-Flanking Sequence of Crp2/SmLim—To determine whether the sequence flanking the 5′ end of the Crp2/SmLim gene had promoter activity, we transfected into RASMC the luciferase plasmid which contains bp 74 to 39 of the mouse Crp2/SmLim gene. Addition of nuclear extracts from RASMC resulted in a retarded DNA-protein complex (arrow on left). The complex was abolished by the addition of an identical unlabeled oligonucleotide (l) but not by the addition of unrelated ATF. An Sp1 consensus sequence oligonucleotide abolished binding. Incubation of nuclear extracts with Sp1 antibodies (AB) (30 ng) before the reaction produced a supershifted band (arrow on right).

Another possibility, which would contain most of the promoter activity. Sp1 Protein Binding to the Core Promoter of Crp2/SmLim—By using the DNA fragment encoding bp 74 to 39 (which contains two Sp1 sites) as a probe in gel mobility shift analysis, we detected a specific DNA-protein complex when the probe was incubated with nuclear extracts prepared from RASMC (Fig. 5). The DNA-protein complex was specific because a 100-fold molar excess of unlabeled identical competitor, but not an unrelated DNA fragment encoding the ATF sequence, abolished the binding complex. In addition, this DNA-protein complex was competed away by a consensus Sp1 oligonucleotide, indicating the presence of Sp1 binding proteins in the complex. To characterize these complexes further, we performed binding reactions in conjunction with antibodies. Antibodies to Sp1 supershifted the upper part of the DNA-protein complex (Fig. 5, arrow on right). In contrast, antibodies to AP2 had no effect (data not shown). These results suggest the presence of Sp1 proteins in the upper DNA-protein complex. The lower part of the complex may contain other members of the Sp family, as observed for the PDGF-B gene by Liang et al. (39).

Preferential Expression of the ~4855SmLIM-lacZ Transgene in the VSMC of Transgenic Mice—To determine whether the 5-kb Crp2/SmLim 5′-flanking sequence contained VSMC-specific elements, we first transfected luciferase plasmids ~4855SmLIM and ~438SmLIM into different cell types in culture. The ~4855SmLIM (Fig. 6) and ~438SmLIM (data not shown) reporter constructs both directed a high level of luciferase gene expression in HeLa, 3T3, and HepG2 cells. These in vitro data raised the possibility that the SMC-specific element(s) were located in a region more 5′ of bp ~4855, in the introns, or in the 3′ region. Another possibility, which would require testing in vivo, was that proper chromatin structure or
DNA methylation was necessary for tissue-specific expression of CRP2/SmLIM. We therefore generated two independent lines of mice harboring a transgene that encodes the 5-kb Crp2/SmLim 5′-flanking region and the bacterial β-galactosidase reporter gene lacZ (with a nuclear localization signal).

Similar results were obtained in both lines of mice. Experiments in one line (line 1822) are shown in Fig. 7. Whole mount staining at embryonic day 11.5 showed high level β-galactosidase activity throughout the developing vasculature (Fig. 7A, blue). Blue staining for β-galactosidase activity was visible in the outflow tract, dorsal aorta, and small blood vessels, as well as in small populations of cells within the neural tube (Fig. 7, A and B). No reporter gene activity was visible within the somites (Fig. 7B). A higher power magnification (Fig. 7C) showed that reporter gene expression was restricted to the smooth muscle cells underlying the endothelium of the dorsal aorta. Blue staining was also visible in the intersegmental arteries branching from the dorsal aorta (Fig. 7C). These results indicate that the 5′-flanking region of Crp2/SmLim directs a high level of VSMC-specific promoter activity in vivo.

**DISCUSSION**

CRP2/SmLIM, a member of the CRP family of LIM-only proteins, is expressed mainly in VSMC in vivo (14). In addition, CRP2/SmLIM is down-regulated markedly in de-differentiated VSMC both in vivo and in vitro (14). This down-regulation implies a role for the gene in the regulation of VSMC differentiation. To elucidate the biological function of CRP2/SmLIM and the molecular mechanisms regulating its expression, we cloned and characterized the mouse Crp2/SmLim gene. Mouse Crp2/SmLim is a 20-kb single copy gene composed of six exons and five introns (Fig. 1). The genomic organization of mouse Crp2/SmLim is very similar to that of human CRP (CRP1) (19), the first member of the CRP protein family. Both genes contain an exon 1 encoding the 5′-untranslated region and a very large intron 1 (~10 kb). The sizes of exons 2 through 5 are identical in both genes. The positions where introns interrupt the coding portion of the cognate protein are also identical, and each finger of the two LIM domains is preceded by an intron. Despite their similarity in genomic organization, however, CRP2/SmLIM and CRP1 are distinct genes that localize to different human chromosomes (40) (data not shown). Nonetheless, their structural homology raises the possibility that the two genes resulted from gene duplication during evolution.

The Crp2/SmLim promoter is rich in G and C residues and has numerous binding sites for Sp1, a ubiquitous nuclear protein that can initiate transcription from TATA-less promoters (41). Many E box motifs are also found in this promoter region (Fig. 3). E box motifs bind to skeletal muscle-specific, basic helix-loop-helix transcription factors that are important for skeletal muscle-specific gene expression (42, 43). Although the function of E box motifs in smooth muscle cell gene expression is still unclear, E box sequences have been found in genes such as smooth muscle α-actin (44) and smooth muscle myosin heavy chain (45, 46).

There is no consensus MEF2 motif in the 5-kb Crp2/SmLim promoter region. There are, however, two AT-rich MEF2-like sequences within the 5-kb promoter region. MEF2 plays a key role in the regulation of skeletal and cardiac muscle-specific gene transcription and acts synergistically with other transcription factors (47, 48). In Drosophila, for example, D-MEF2 is important for determining the lineage of skeletal, cardiac, and smooth muscle cells (49). One GATA motif in Crp2/SmLim is located immediately 3′ of a potential MEF2-like site (bp −1507). One C-rich motif (bp −1628) immediately 5′ of a
CAAT motif is identical to the CCAC box in the human myoglobin and troponin Ic genes. A CCAC box motif, in conjunction with an AT-rich motif, is essential for full activity of the myoglobin promoter in myocardium (37).

A consensus GArC motif, important for regulating expression of the human cardiac myosin heavy chain gene (36), is located at Crp2/SmLim bp −3103. Recent studies of the smooth muscle α-actin, myosin heavy chain, and SM22α promoters have shown that CArG elements are essential for smooth muscle-specific expression (44, 46, 50, 51). Surprisingly, there is no consensus CARG box sequence within the 5-kb region of the Crp2/SmLim 5′-flanking sequence.

The 5-kb Crp2/SmLim 5′-flanking sequence directed a high level of transcriptional activity in VSMC (Fig. 4), consistent with the finding that CRP2/SmLim is expressed preferentially in this cell type. However, this DNA fragment also directed reporter gene expression in non-smooth muscle cell types such as HeLa, 3T3, and HepG2 at comparable levels (Fig. 6). This in vitro finding is similar to observations about other promoters of cell type-specific genes, such as MyoD (52) and SM22α (53), whose expression is much more restricted in vivo than in vitro.

It is possible that proper chromatin structure or DNA methylation (54–58) is required for tissue-specific expression of Crp2/SmLim in vivo. To determine if this 5-kb of Crp2/SmLim 5′-flanking region contained VSMC-specific cis-acting element(s), we generated transgenic mice harboring this DNA fragment fused to the bacterial β-galactosidase reporter gene lacZ. The reporter gene construct was expressed mainly in the VSMC of two independent lines of transgenic mice (Fig. 7 and data not shown). These observations suggest that the 5-kb Crp2/SmLim 5′-flanking sequence contains regulatory element(s) necessary for its expression in the VSMC of transgenic mice. Furthermore, the pattern of transgene expression was very similar to that of endogenous gene expression.2 The lacZ reporter gene was also expressed in VSMC of both the aorta and the small arteries (Fig. 7A). This pattern differs from that of the SM22α promoter, which is expressed preferentially in the large arteries (51, 53).

One of the hallmarks of arteriosclerosis is VSMC proliferation. Very little is known, however, about nodal genes regulating the growth and differentiation of this cell type (59). The Crp2/SmLim gene and 5′-flanking sequence are important tools that will allow us to investigate the molecular mechanisms regulating VSMC growth and differentiation. Moreover, knowledge about the cis-acting elements that restrict gene expression to VSMC will be useful for directing specific gene products to this cell type.

Acknowledgments—We are grateful to the late Edgar Haber for editorial assistance.

REFERENCES
