Cloning and analysis of the promoter region of the rat SM22α gene

Paul R. KEMP,* Jane K. OSBOURN, David J. GRAINGER and James C. METCALFE
Department of Biochemistry, Cambridge University, Tennis Court Road, Cambridge CB2 1QW, U.K.

We have cloned and sequenced a 1.9 kb fragment of the 5′-upstream sequence of the smooth-muscle-specific gene SM22α. The region cloned consisted of the SM22α promoter, a 65 bp exon containing most of the 5′-untranslated region and 307 bp of the first intron. A 1.5 kb fragment at the 5′ end of this sequence was able to drive the expression of a reporter chloramphenicol acetyltransferase (CAT) gene in both vascular smooth-muscle cells and Rat-1 fibroblasts. This promoter region did not contain a consensus TATAA box but contained the sequence TTTAAA 25 bp from the major start site identified by primer extension. Deletion analysis showed that a fragment of the promoter from +65 to −303 was more active in both cell types than the 1.5 kb fragment suggesting that there are silencer sequences in the region 5′ to the core promoter. CAT activity was also observed with fragments containing bases +65 to −193 and +65 to −117 in smooth-muscle cells. In contrast with the smooth-muscle cells, no CAT activity was detected in Rat-1 fibroblasts with the smallest two fragments. The residual promoter activity in the smallest fragment of the SM22α promoter tested suggested that, unlike the smooth-muscle α-actin promoter, transcription from the SM22α promoter can occur in smooth-muscle cells in the absence of factors binding to $CC(A/T_{rich})_6GG$ (CArG box) or CANNTG (E box) motifs.

INTRODUCTION

The differentiated state of vascular smooth-muscle cells (VSMCs) is marked by the presence of smooth-muscle-specific isoforms of several structural proteins [e.g. smooth-muscle myosin heavy chain (SM-MHC) and smooth-muscle α-actin (SM α-actin)] and their regulatory proteins (e.g. myosin light chain kinase and calponin). During aortic development the mesenchymal cells surrounding the aortic endothelium express such smooth-muscle markers although the factors that control this process remain unknown [1]. To investigate the process of smooth-muscle differentiation the genomic structures of two genes (SM α-actin and SM-MHC) have previously been determined. The most extensively studied of these genes is SM α-actin for which the promoter region has been cloned from rat [2], mouse [3], human [4] and chicken [5] DNA. The rat, mouse and human promoters contain three conserved CArG boxes and a conserved E box within 300 bp of the transcription start site. Deletion studies of the SM α-actin promoter have led to the conclusion that CArG box-like elements play an important role in the expression of this gene [2]. However, CArG boxes also play a role in the expression of many non-smooth-muscle-specific genes through the activity of the transcription factor SRF [6–8], and this factor has also been shown to drive SM α-actin gene expression in non-smooth-muscle cells (e.g. Ras-transformed fibroblasts [8]). These observations suggest that, although CArG boxes are important in determining the level of SM α-actin gene expression, they do not appear to be the elements that determine the tissue-restricted expression of this gene. Thus SRF is unlikely to play a defining role in smooth-muscle differentiation. CArG boxes are also present in the promoter region of rabbit SM-MHC together with six E boxes and two MEF-2-like binding sites [9]. Deletion analysis of this promoter has implicated a region containing the most distal MEF-2-binding site and the region between −188 and −509 as positive regulatory regions. However, the factors that drive smooth-muscle-restricted expression of these genes remain obscure and analysis of other smooth-muscle-specific promoters should help in their identification.

In the embryo, SM α-actin together with several other proteins that are predominantly or exclusively found in adult smooth muscle (e.g. SM-MHC and desmin) can be found in non-smooth-muscle tissues [1]. However, in avian systems at least, the expression of other genes (e.g. SM22α, calponin and myosin light chain kinase) appears to be restricted to smooth muscle in the embryo as well as in the adult [1]. These genes may therefore provide a better system for analysis of smooth-muscle-specific gene expression. SM22α is a protein originally isolated from chicken gizzard [10] with extensive similarity to the actin-binding proteins transgelin [11] and a 25 kDa protein from bovine smooth muscle [12]. The gene that codes for SM22α is expressed at high levels in differentiated smooth-muscle tissue but does not appear to be expressed in other tissues [13]. In avian systems the protein appears in the aorta after 4 days of development (E4), before the appearance of calponin at E6 but after the appearance of myosin light chain kinase, SM α-actin and desmin [1]. The appearance of SM22α and calponin has therefore been suggested to mark the late stage of smooth-muscle differentiation. *In vitro*, expression of SM22α is highest in freshly dispersed VSMCs that are highly differentiated (as marked by their expression of SM-MHC) and much lower in dedifferentiated VSMCs that have been passaged for many generations. Chicken [14], rat [15,13] and human [16] SM22α cDNAs have been cloned, and the sequences of these clones demonstrated marked identity with the smooth-muscle protein, calponin [17]. Northern-blot analysis of rat aortic VSMCs has shown two transcripts of SM22α [13] and both have been cloned. The shorter of these transcripts (999 bp [13]) appears to be extended by the addition of 114 bp to the 5′-untranslated leader and 73 bp to the 3′-untranslated sequence to generate the longer transcript (1186 bp) described by Nishida et al. [15].

In this study we have cloned and sequenced the 5′ upstream region of the SM22α gene. We have used transient transfection of deletion mutants of this promoter fused to chloramphenicol

Abbreviations used: SM-MHC, smooth-muscle myosin heavy chain; VSMC, vascular smooth-muscle cell; DMEM, Dulbecco's modified Eagle's medium; FCS, foetal calf serum; CAT, chloramphenicol acetyltransferase; AMV-RT, avian myelomavirus reverse transcriptase; RT, reverse transcription.

* To whom correspondence should be addressed.

acetyltransferase (CAT) as a reporter gene to delineate the regions of the promoter responsible for the strong expression of SM22α and its smooth-muscle-restricted expression.

MATERIALS AND METHODS

Reverse transcription (RT)-PCR and PCR

Cytoplasmic RNA (2 μg) isolated as described previously [18] was denatured in a total volume of 11 μl at 65 °C for 10 min and then chilled on ice. To this was added 4 μl of 5 × avian myelomavirus reverse transcriptase (AMV-RT) buffer (Promega), 2 μl of 2 mM dNTPs, 1 μl of 100 mM dithiothreitol, 0.5 μl of RNAsin (Promega) and 0.5 μl of primer (5 pmol/μl). The reaction mixture was incubated at 42 °C for 1 h before 0.5 μl (2 units) of AMV-RT was added. After a further 45 min at 42 °C the reaction was terminated by placing the mixture on ice. PCR amplification was carried out by adding 5 μl of the cDNA to a mixture containing 4 μl of 10 × reaction mix (Perkin–Elmer), 5 μl of 2 mM dNTPs, 2.5 μl of primer 1 (5 pmol/μl), 4 μl of 25 mM $MgCl_2$ (to give a final concentration of 2 mM) and 20 μl of water. The reaction was started by heating to 95 °C for 30 s, cooling to 80 °C and adding 10 μl containing 1 μl of 10 × reaction mix (Perkin–Elmer), 2.5 μl of primer 2 (0.5 pmol/μl) and 0.5 μl of *Taq* polymerase (2.5 units; Perkin–Elmer). The reactions were performed using 30 cycles of 1 min at 95 °C, 1 min at 64 °C, 1 min at 72 °C. The amplification products were purified from low-gelling-temperature agarose, phosphorylated with T4 polynucleotide kinase, made blunt-ended with T4 DNA polymerase and then cloned into Bluescript KS+.

Plasmid DNA was linearized with *SalI* and resuspended at 10 ng/μl before amplification for 30 cycles as described above except that the first denaturation cycle was extended to 4 min. Rat genomic DNA was isolated from cultured VSMCs. Cells grown to confluence were washed twice with PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4) and harvested by scraping into 0.5 ml of ice-cold lysis buffer (150 mM NaCl, 10 mM Tris/HCl, pH 7.4, 1 mM $MgCl_2$ and 0.5% Nonidet P40). The sample was incubated on ice for 5 min and then spun in a microcentrifuge for 5 min. The pellet was resuspended in NDB (75 mM NaCl, 24 mM EDTA, pH 8) plus 100 μg/ml proteinase K and 0.5% SDS then incubated overnight at 37 °C. After extraction with phenol (three times) and chloroform (once), the DNA was precipitated with ethanol. Then 2 μg of DNA was amplified in the reaction mixture described using 35 cycles of 95 °C for 1 min, 63 °C for 1.5 min, 72 °C for 4 min except that the first denaturation step was extended to 4 min and the last extension step was extended to 14 min.

Library screening and Southern-blot analysis

λ clones containing the SM22α 5′-flanking region were isolated by screening 1×10^6 plaques of a rat genomic library in λGEM11 (Promega) with a ^{32}P-labelled probe to the first intron of SM22α. Plaques were lifted on to nylon membranes (Hybond N), the DNA denatured as described by Maniatis et al. [19] and cross-linked to the filter by UV irradiation. Filters were hybridized to the probe at 65 °C in 3 × SSC (1 × SSC contains 150 mM NaCl and 15 mM sodium citrate), dextran sulphate (5%), 10 × Denhardt's solution (1 × Denhardt's solution contains 0.02% Ficoll 400, 0.02% polyvinylpyrrolidone and 0.02% BSA), salmon sperm DNA (250 μg/ml) and SDS (0.1%) overnight. The filters were washed twice at 65 °C for 1 h in 1 × SSC/0.1% SDS before exposure to Fuji RX X-ray film.

Southern-blot analysis [19] was performed on purified λ DNA from the positive clones using both the cDNA and genomic probes. DNA, UV-cross-linked to nylon membranes, was hybridized to the probes and washed as described for the phage lifts. Southern blots were analysed using a Molecular Dynamics 425 phosphorimager.

Primer extension

Polyadenylated RNA was prepared from 25 μg of total RNA using the Qiagen direct mRNA kit. The RNA was then reverse-transcribed as described for RT-PCR in the presence of 10 pmol of primer labelled with [γ-^{32}P]ATP as described by Maniatis et al. [19]. The reaction product was extracted with phenol/chloroform, precipitated with ethanol and the products were run on a 6% acrylamide sequencing gel.

Cell culture, transfection, CAT assay and Northern-blot analysis

Subcultured VSMCs from adult rat aortae were derived as described previously [20], grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with penicillin and streptomycin plus 10% foetal calf serum (FCS) and passaged 1:3 every 4 days by trypsin treatment. Rat-1 fibroblasts were cultured in DMEM supplemented with penicillin and streptomycin plus 10% FCS and passaged 1:5 every 3 days. Cells were transfected by electroporation at room temperature as previously described [21] at 280 V, 960 μF using 30 μg of test plasmid DNA and 10 μg of pSVβgal (Promega) DNA. After electroporation, the cells were allowed to recover for 10 min at room temperature before being added to DMEM supplemented with penicillin and streptomycin plus 10% FCS (9.2 ml) and plated into 100 mm dishes. After transfection, cells were incubated for 24 h before being harvested for CAT assay by TLC and β-galactosidase assay. These procedures were carried out as described by Maniatis et al. [19]. Northern-blot analysis was carried out as described previously [18] using a cDNA probe encoding bases 9–487 of SM22α.

RESULTS

Generation of probes for SM22α

Initial attempts to isolate a cDNA probe for SM22α by RT-PCR used a primer to the 5′ end of the 1186 bp sequence, the longer of the two transcripts [15] (primer A, Table 1) and a primer to the 5′ end of the coding sequence (primer B, Table 1). This reaction failed to produce the expected product (a 286 bp fragment), and analysis of primer A showed that it was the reverse and complement of bases 389–415 of the 1186 bp sequence. This complementary sequence implied the potential for considerable

Table 1 Oligonucleotide primers used for PCR and primer

Primer	Sequence
Primer A	AGGGCGGGTTCTCAGGCACCTTCACTG
Primer B	CATCAGGGCCACACTGCATTACAATCC
Primer C	CGGTAGTGTCCATCGTTCTTGGTCAC
Primer D	TCCTTCCAGCCCACAAACGACCAAGC
Primer E	CAATCCACTCCACTAGCCGCTCCTC
Primer F	CCGCTCCTCCAGCTCCTCATCATAC
Primer G	AAGGCTTGGTCGTTTGTGGGCTGGAAG
Primer H	GGAAGGTTTTCGTGGTCCTGCC
Primer I	CCGACAGACTGCTCCAACTTGGTG
Primer J	TGGGGTGGTGTGGAAGCCAAGC

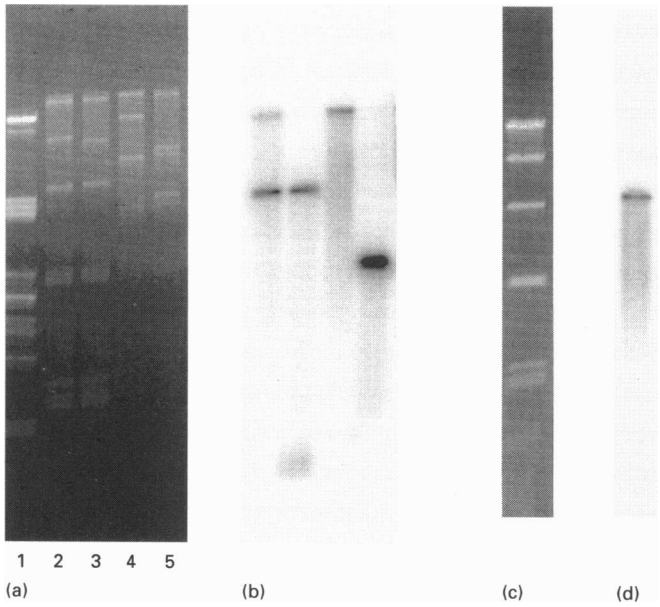

Figure 1 Southern-blot analysis of a λ clone containing the 5′ end of the SM22α gene

(**a**) Ethidium bromide-stained gel of λ *Pst*I markers (lane 1) and λ16 digested with *Bam*HI, *Bam*HI–*Sal*I, *Hin*dIII and *Bam*HI–*Hin*dIII (lanes 2–5 respectively). (**b**) Southern blot of gel shown in (**a**) probed with cDNA probe. (**c**) Ethidium bromide-stained gel of λ16 digested with *Bam*HI. (**d**) Southern blot of gel shown in (**c**) probed with genomic DNA probe.

RNA secondary structure at the 5′ end of the long transcript. Further analysis of the structure of the mRNA using the RNA structure program FOLDRNA [22] showed that bases 1–106 of the sequence described by Nishida et al. [15] were the exact reverse and complement of bases 310–415 of the same sequence. The secondary structure formed by the complementary sequences may have prevented the initiation of reverse transcription in the RT-PCR experiment. Alternatively some or all of the complementary sequence at the 5′ end may be a sequencing artifact. A second set of PCR primers (primers C and D) was therefore used to generate a cDNA probe corresponding to the 5′ end of the shorter published sequence (bases 9–487 [13]) and including the region complementary to the 106 bp at the 5′ end of the sequence published by Nishida et al. (15). The cDNA produced by this reaction was cloned into pBluescript and sequenced. In order to generate a genomic probe, primers B and C were used to amplify a 3.6 kb intron of the gene [23] by PCR from genomic DNA isolated from rat aortic VSMCs. The DNA generated by this reaction was cloned into pBluescript and the identity of this clone was confirmed by the sequencing of both ends of the insert.

Genomic mapping and identification of the 5′ end of the gene

To establish the genomic structure of the SM22α gene, λ clones were isolated from the genomic library by screening with the genomic probe described above. These clones were mapped by restriction mapping and Southern blotting using either the cDNA probe or the genomic probe. Southern-blot analysis of one clone (λ16) showed that both the cDNA and the genomic probes hybridized to a 5.2 kb *Bam*HI fragment with the cDNA probe also hybridizing to rat genomic sequence attached to the λ left arm (Figure 1). A *Bam*HI–*Sal*I double digest showed that approximately 1 kb of the SM22α gene remained attached to the λ left arm in the digest which used *Bam*HI alone. Digestion of the clone with *Bam*HI and *Hin*dIII resulted in cleavage of the 5.2 kb *Bam*HI fragment into a 3.2 kb fragment and a 2.0 kb fragment. Only the smaller of these two fragments hybridized to the cDNA probe. Since the cDNA probe contained the sequence complementary to the 106 bp at the 5′ end of the sequence published by Nishida et al. [15], the absence of hybridization by this probe to the 3.2 kb band in the *Bam*HI–*Hin*dIII double digest or to any additional *Bam*HI fragment suggested that some or all of bases 1–106 had arisen through a sequencing artifact. Further restriction mapping gave the restriction and genomic maps shown in Figure 2.

The size of the 5′-untranslated region of the mRNA was determined by primer extension. Polyadenylated RNA was isolated from VSMCs and Rat-1 fibroblasts and reverse-transcribed from two separate end-labelled primers E and F. Primers E and F were designed so that transcription start sites

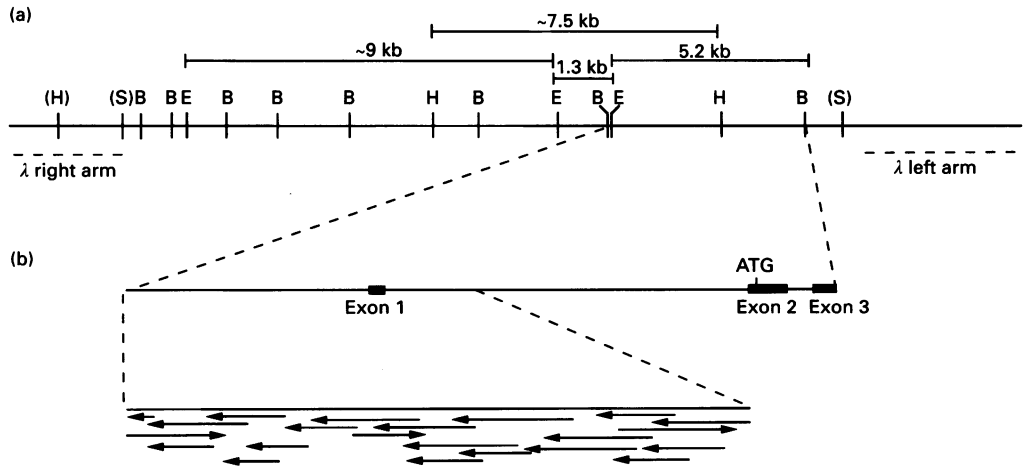

Figure 2 Physical map of the 5′ end of the rat SM22α gene

(**a**) Restriction map of recombinant bacteriophage λ16. Restriction enzyme sites are shown: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; S, *Sal*I. Restriction sites given in parentheses are in the λ arms. (**b**) Physical map of the 5.2 kb *Bam*HI fragment. Exons are represented by thick lines. The sequencing strategy for the 5′ end of the gene is shown by the arrows underneath.

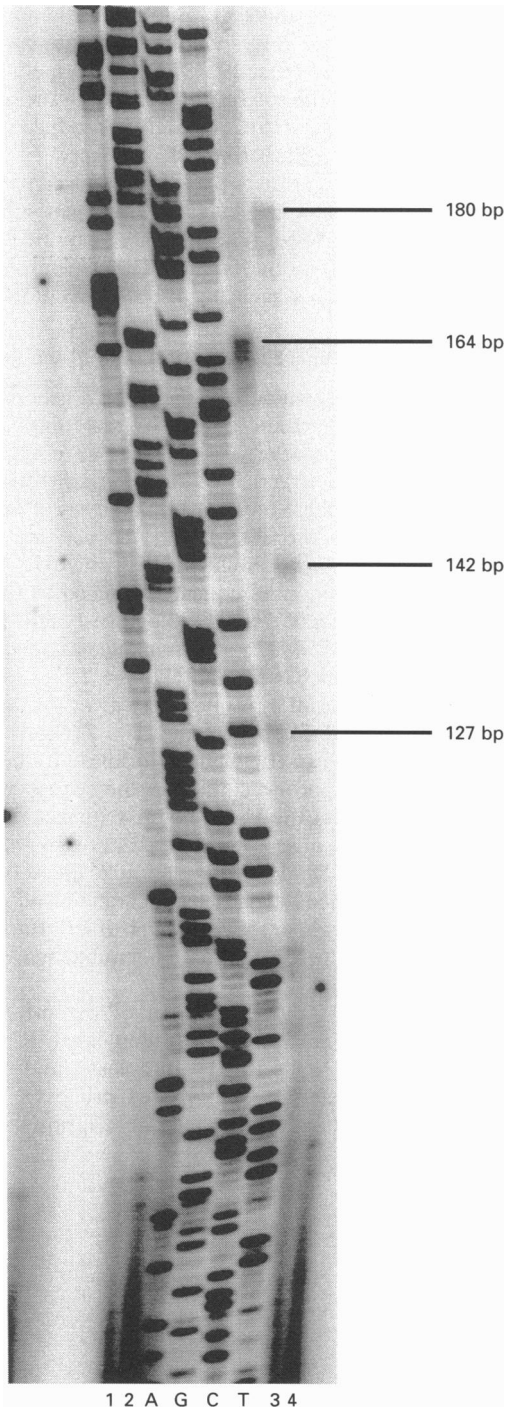

Figure 3 Primer-extension analysis of the rat SM22α gene

Polyadenylated RNA from Rat 1 fibroblasts (lanes 1 and 2) or VSMCs (passage 11; lanes 3 and 4) was reverse-transcribed from ^{32}P-end-labelled primer E (lanes 1 and 3) or ^{32}P-end-labelled primer F (lanes 2 and 4). Sizes of extension products were determined by comparison with a DNA sequencing ladder of known length (lanes A, G, C, T).

```
GGATCCCGTC TAGTGAATTC AGGACATAAT CAGTGGCTGG GAGTCCCCTG
                                                    -1416
GACGGCAAGT GCCATCTCTG GGGAACTCCA GCTTCCTGAT TAATGACCCT

TCCTTCAGAT GCCACGAGGA GGTGCTGTAG CTCTATGCAC CAGTAACTTA
                                                    -1316
AATCAGGCGG CCAGTTCTGG GTGAGACTGA CCCTGCCTGA AGAGCCCTTT

CTTCCTCCTA CTCCTTTCTC CTTCTCCCTC TATCAGGTCC CTGAAGTTCC
                                                    -1216
CAGAATTGGG GGCAGAGCTC AGAGACACCA CTAAAGTCTT ACCTTTTAAA

AAGTGACATT TAGTGAGTGA GACAGACAGC AGCATGTGCA GGTCAGAGGA
                                                    -1116
GAGCCGGTTC TGTCAATACT GTGTGTGTTC TGGGTATTGA ACTCAGACCA

TCAGGTTTGA CAGCAGTTGT CTAACCCGTG AACCTGTCCC TTCCCAGGCC
                                                    -1016
CACTGTAGCT AGGTACAAAG GGGATAAGTG AGAACCTCTA AGGGGTAGAA

TCTTTCAAGA AAAGGACTGT TTGGAAGGGA AGGGAGGGGT GACCATGTTT
                                                    -916
CCTAGTCCTC TTTCTTTCTT TTTCTTTTTT TTTTTTAAGA TTTATTCATT
     MEF-2-like
TACTATATTA TAAGTACACT GTAGCTGTCT TCAGACACAC CAGAAGAGGG
                                                    -816
GATCAGATCT CTTTACAGAT GGTTGTGAGC CACCATGTGG TTGCTGGGAA

ATGAACTCAG GACCTCTGGA AGAGCAGACA GTGCTCTTAA CCACTGAACC
                                                    -716
ATCTCTCCAG CCCCCTAGTC CTCTTTCTTA CCCTTGTGTC TTCAGCTTTG

CCAGGCCTTA AACATCCTCC CGTTGTCACT GTGCTCTAGA TGGGGCCAGG
                                                    -616
GTTGACTTGC CGCTGAGCAA GTCACCCCCT AGTGAAGCTC CAAGTAGAAC

TGTCCCCTAC CAGTGGGCAG TCATACCCAA GTTCTACCAC TCTGAACCTC
                                                    -516
AGTTTTCACA GCTGTCAAAG GGGCTGGTAA CTCCTACCTC TTAACAGCTC

TGAGTGTGAA AGGAAGGCCT TTGCTTAGTG CCTGGTGTGC AGCCAGGCTG
                                                    -416
CAGTGGAGAC TAGTTTCCAC CAACTGGATT TTAAAGCCTT TCAAAAAGGT

GGTTTGTTTG TCCCCTGCAA GCTCCTAAGA CGTTTCTCTG TCGGGCCAAA
                                                    -316
TTCTCTAGAG CACCTCCTCC TTTCTTTCTC ATTGAAAAGC AGACCCAAGT
                                        CArG-box    CArG-box
CGGGGCAGCA AGGAAGGTTT TCGTGGTCCT GCCCATAAAA GGTTTTTCCC
                  GC box                            -216
GCCACCCTCA GCACCGCCCC GCCCCAACCC CGCAGCATCT CCAAAGCATG

CAGAGAATGT CTCGGGCTGC CCCCGACAGA CTGCTCCAAC TTGGTGTCTT
      CArG-box                          E-box  GC box-116
TCCCCAAATA TGGAGCCTGT GTGGAGTGAG TGGGGCAGCT GGGGCGGGTG

GGGTGGTGTG GGAAGCCAAG CAGACTTCCA TGGGCAGGAG GGGCGCCAGT
                                                    -16
GGGCAGGCAG AGAGGTGACA TCACTGCCTC GGAGGCCTTT AAACCCCTCA
                  +1
CCCAGCCGGC  GCCCCAGCCC GTCTGCACCA GCCCAGACAG CAACGCTACT
                                                    +85
CTCCTTCCAG CCCACAAACG ACCAAGCCTT GTAAGTGGGA GCAACGGGAG

CAGCAGGGTT GTGTGTTATT AGGCAACGCT GGGGGTTCAG CTGGGGGGTC
                                                    +185
CTCCAGTCTC CTCTAGCTTG CTGGGTGGAA TTGGGTCTCT GAGTAGGCAA

AAAGTCTGGT TAAAATGCCT AGGGTGACTT GACTGGGGCA GAAGAATCAA
                                                    +285
ATGGGGTGAG AGGGAGGCTT GCATTTCAGT CACTCGCCCA CCTATGGGCT

GCGTGTTAGA CCCAAAGAGG TGCTGAATAA GGCCAAGGGA CCAACTCCAT
                                        +374
CTCAGCATCT CCAAGATGAG GGTCTGAGTT CTAGAAAGC
```

Figure 4 Sequence of the promoter and 5′ end of rat SM22α

Nucleotide sequence of the 5′ region of the rat SM22α gene. Exonic sequence is shown in bold and regions with similarity to some of the response elements described in the text are underlined and identified.

would be seen as a band in the primer E lane and a similar band 16 bases larger in the primer F lane. Extension from primer E resulted in one product of 127 bases and a group of three bands of 162, 163 and 164 bases (Figure 3). Extension from primer F resulted in one product of 142 bases and a group of three bands of 178, 179 and 180 bases (Figure 3). The group of three products produced by the two primers are separated by 16 bases, suggesting that they are genuine start sites. However, the 142-base product

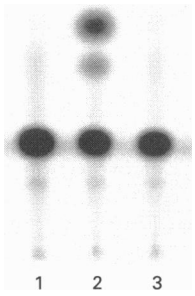

Figure 5 CAT activity in transiently transfected VSMCs

Promoter constructs containing the 5′-upstream region of the SM22α gene fused in either orientation to the CAT gene were generated. The activity of these constructs was tested in VSMCs by transient transfection as described in the Materials and methods section. Lane 1, pCB53; lane 2, pCB55; lane 3, pCAT-basic. After 24 h the cells were harvested, cell extracts prepared and CAT activity was assayed.

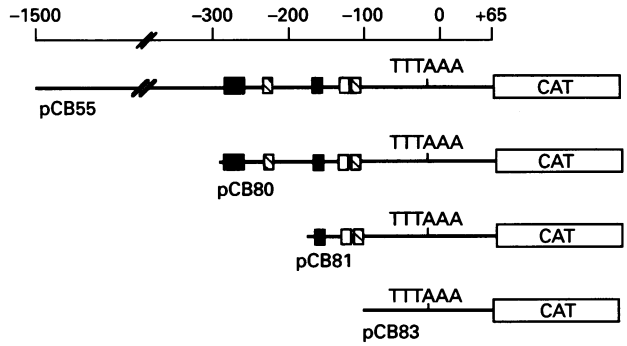

Figure 6 SM22α promoter constructs

Fragments of the SM22α promoter were generated by PCR using primers designed to remove potential transcription-factor-binding sites. These fragments were then fused to the CAT gene and used in transient transfection assays of Rat-1 fibroblasts and VSMCs. ■, CArG box; □, E box; ▨, GC box.

produced by extension of primer F is only 15 bases larger than the 127-base product produced from primer E and may not therefore represent a genuine start site for SM22α. No primer extension occurred in RNA isolated from the Rat-1 fibroblasts suggesting that the products observed are due to the presence of significant amounts of SM22α mRNA in the smooth-muscle cells. The data show that the start site for SM22α is 28 bp upstream of the end of the sequence described by Shanahan et al. [13], and in Figure 4 this position is marked as +1.

Sequencing of the promoter region of the SM22α gene

The 5.2 kb *Bam*HI fragment from λ16 was cloned into *Bam*HI-cut pBluescript KS+ and the orientation of the insert was determined by restriction mapping and confirmed by sequencing using primers made to the T7 and T3 promoters. The sequence of 1.892 kb at the 5′ end of this clone was determined by dideoxy sequencing of exonuclease III deletions and is shown in Figure 4. The sequence included 303 bp of intron, 37 bp of the 5′ end of the SM22α cDNA described by Shanahan et al. [13] and a further 14 bp of the sequence described by Nishida et al. [15]. This observation indicates that 99 bp at the 5′ end of the cDNA sequence described by Nishida et al. [15] is an artifact. The sequence obtained did not contain a classical TATAA box, but did contain three CArG boxes (−162, −275, −283), an E box (−130), two GC boxes (−123 and −252), three AP2-binding sites (−717, −253 and −193), an NFκB site (−1478), an AP1 site (−1430) and a C/EBP site (−1311). An AT-rich sequence similar to the MEF-2-like site in rabbit SM-MHC [9] was also present at −912.

Determination of the promoter activity

To demonstrate promoter activity in the absence of any potential regulatory sequences in the 3.6 kb intron, the promoter region was isolated by PCR between primer G (Table 1) and the T7 sequencing primer. This fragment was cloned by blunt-end ligation into pCAT-basic. Vectors corresponding to the promoter in the correct (pCB55) and opposite (pCB53) orientations were generated and transfected by electroporation into rat aortic VSMCs. After 24 h of incubation, cells transfected with pCB55 showed a high level of CAT activity, whereas cells transfected with either pCB53 or pCAT-basic alone showed no detectable CAT activity (Figure 5).

Table 2 CAT activity in transiently transfected VSMCs and Rat-1 fibroblasts

VSMCs and Rat-1 fibroblasts were transfected with 30 μg of the promoter–CAT fusion constructs described in Figure 6 and 10 μg of the β-galactosidase expression vector pSVβgal, plated into 90 mm dishes and cultured in DMEM + 10% FCS. Cell extracts were prepared 24 h after transfection and used for CAT assay and β-galactosidase assay. CAT activity was normalized to the β-galactosidase activity of the extract and is expressed as percentage of the activity in pCAT-control (the CAT gene under the control of a simian virus 40 promoter with a simian virus 40 enhancer; Promega). Results are means ± S.E.M. ($n = 3$ separate transfections).

Vector	CAT activity in VSMCs (%)	CAT activity in Rat-1 fibroblasts (%)
pCAT-control	100	100
pCAT-Basic	4 ± 1	2 ± 1
pCB55	39 ± 5	17 ± 5
pCB80	294 ± 45	280 ± 28
pCB81	44 ± 17	3 ± 1
pCB83	19 ± 5	2 ± 1

A series of fragments was then generated by PCR to remove the potential transcription-factor-binding sites identified from the sequence. PCR between primer G and primers H, I or J generated fragments of 373, 252 and 180 bases respectively. These fragments were cloned into pCAT-basic to give the constructs shown in Figure 6. The sequence and orientation of each fragment was checked by sequencing and the ability of each construct to drive expression of the CAT gene was tested in VSMCs and Rat-1 fibroblasts by transient transfection and CAT assay. Construct pCB80 was active in both cell types and showed greater promoter activity than any other construct tested (294 ± 45% and 280 ± 28% of pCAT-control in VSMCs and Rat-1 fibroblasts respectively; Table 2). Removal of the SP1 site and the two most 5′ CArG boxes (pCB81) reduced CAT activity to 44 ± 17% of pCAT-control in VSMCs and abolished CAT activity in Rat-1 fibroblasts (Table 2). The smallest construct tested (pCB83) had 19 ± 5% of the activity of pCAT-control and this activity was also restricted to VSMCs (Table 2).

Since both the Rat-1 fibroblasts and the VSMCs showed

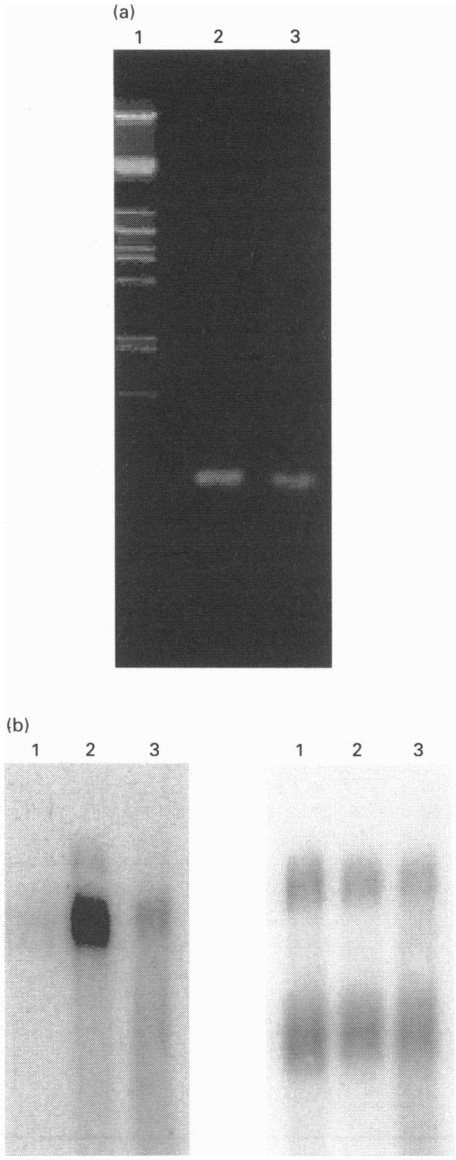

Figure 7 Northern-blot and PCR analysis of SM22α mRNA from VSMCs and Rat-1 fibroblasts

Passage-11 VSMCs, passage-30 VSMCs and Rat-1 fibroblasts were seeded into 90 mm dishes and grown for 3 days in DMEM + 10% FCS before RNA was extracted. (**a**) 2 μg from each sample was reverse-transcribed from primer C then amplified through 30 cycles of PCR using primers C and D as described in the Materials and methods section. Lane 1, λ digested with *PstI*; lane 2, RNA extracted from passage-30 VSMCs; lane 3, RNA extracted from Rat-1 fibroblasts. (**b**) 10 μg from each sample was loaded on to a 1.5% agarose/formaldehyde gel. After transfer to nitrocellulose the RNA was hybridized to a radiolabelled SM22α cDNA. After exposure to a phosphor screen for 24 h the probe was removed by boiling in 0.1% SDS and the blot reprobed for glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Lane 1, RNA extracted from Rat-1 fibroblasts; lane 2, RNA extracted from passage-11 VSMCs; lane 3, RNA extracted from passage-30 VSMCs.

similar CAT activity from the longer promoter fragments, the relative level of expression of SM22α was compared between the two cell types by RT-PCR and Northern-blot analysis (Figure 7). RT-PCR confirmed that SM22α was expressed in both cell types (Figure 7a). However, Northern-blot analysis showed that, although a high level of expression of SM22α was found in the VSMCs, the transcript in the Rat-1 fibroblasts was barely detectable (Figure 7b).

DISCUSSION

To examine the control of expression of the smooth-muscle-specific gene SM22α we have isolated and studied the 5′ upstream region of this gene. Identification of the transcription start site and Southern blotting suggested that 99 bp of a previously published sequence [15] was either a sequencing artifact or separated from the rest of the gene by more than 14 kb. As a result, the two transcripts that have been seen for SM22α ([13] and Figure 7b) may differ only by the 3′ extension described by Nishida et al. [15].

The sequence of the SM22α promoter was not found to contain a consensus TATA box. Instead the sequence TTTAAA situated between −23 and −29, which is similar to the TATA box in the cardiac troponin C gene (TTAAA [24]), may function as the TATA box for the SM22α gene. The SM22α promoter possessed multiple CArG boxes, like the smooth-muscle-specific promoters of the SM α-actin gene [2–4] and SM-MHC [9]. Two of the CArG boxes overlap in the SM22α promoter (at −275 and −283) and are adjacent to an SP1-binding site (−252). Removal of the region containing these motifs causes a reduction in the level of CAT activity by approximately 7-fold indicating that one or more of these binding sites is important in the high level of SM22α expression in smooth-muscle cells. Removal of the third CArG box halved the remaining CAT activity in the smooth-muscle cells to 19% of the level of activity from pCAT-control. These data indicate that, unlike the SM α-actin promoter, the SM22α promoter is active in the absence of SRF binding to CArG boxes.

In addition to the multiple CArG boxes, the SM22α promoter contains two GC boxes (−123 and −252) and an E box (−130). Some GC boxes have been shown to bind the transcription factor SP1 in other promoters (e.g. cardiac α-actin), and members of the basic helix–loop–helix (bHLH) family of transcription factors acting through E boxes [25,26] have been shown to play an important role in the expression of skeletal-muscle-specific genes. The residual CAT activity measured in VSMCs but not in Rat-1 fibroblasts after transfection with pCB83 suggests that the E box at −130 is not required for smooth-muscle-specific expression from the SM22α promoter. These data are in agreement with other circumstantial evidence that bHLH transcription factors do not play a significant role in transcription from smooth-muscle-specific promoters since the Id gene (a negative regulator of bHLH transcription factors) can be expressed at high levels in VSMCs at the same time as smooth-muscle-specific genes [18].

Both VSMCs and Rat-1 fibroblasts generated CAT activity from a CAT gene under the control of 1.57 kb of SM22α upstream sequence, although a higher level of CAT activity was found in the VSMCs. Removal of bases −303 to −1565 increased CAT activity in both cell types, suggesting that there are silencer sequences in the region between these bases. Removal of bases −193 to −303 reduced CAT activity in the VSMCs by approximately 7-fold, indicating that SP1 or another transcription factor binding to this region is responsible for a large part of the promoter activity in smooth-muscle cells but is not essential for some expression to occur. Removal of the same region abolished CAT activity in the Rat-1 fibroblasts, showing that the binding of SP1 or other factors to this region is a prerequisite for SM22α promoter-driven CAT activity in these non-smooth-muscle cells. It also suggests that factors binding to the proximal 193 bases of

the SM22α promoter play an important role in restricting expression of SM22α to smooth-muscle cells.

In vivo, the expression of SM22α has been found only in smooth-muscle tissues [13]. In cell culture, however, a low level of SM22α expression has been reported for Rat-2 fibroblasts [13] and senescent human fibroblasts [16] as well as the expression in Rat-1 fibroblasts reported here. As a result some expression of the reporter gene in the Rat-1 fibroblasts is not unexpected. The high level of CAT activity observed from the constructs pCB55 and pCB80 when compared with the low level of mRNA for SM22α in the cells may occur because the promoter DNA in the plasmid constructs used is unmethylated at the CpG sites. Such CpG methylation has previously been shown to prevent expression from the promoter of rat glia-derived nexin in non-glial cells but to allow expression in glioma cells [27]. CpG methylation has also been found to play an important role in the inactivation of DNA *in vivo* and plays a role in X chromosome inactivation [28].

In conclusion we have cloned and sequenced 1.5 kb of the promoter region of a smooth-muscle-specific gene, SM22α. This region contains both positive and negative acting regulatory elements. Transient transfection studies showed that fragments longer than 303 bp were active in non-smooth-muscle cells as well as VSMCs but expression from the proximal 193 bp was restricted to VSMCs.

This work was supported by a BHF group grant to J.C.M. and Dr. P. L. Weissberg. D.J.G. was the recipient of a Wellcome Prize Fellowship.

REFERENCES

1 Duband, J. L., Gimona, M., Scatena, M., Sartore, S. and Small, J. V. (1993) Differentiation **55**, 1–11
2 Blank, R. S., McQuinn, T. C., Yin, K. C. et al. (1992) J. Biol. Chem. **267**, 984–989
3 Min, B., Foster, D. N. and Strauch, A. R. (1990) J. Biol. Chem. **265**, 16667–16675
4 Reddy, S., Ozgur, K., Lu, M. et al. (1990) J. Biol. Chem. **265**, 1683–1687
5 Carrol, S. L., Bergsma, D. J. and Schwartz, R. J. (1986) J. Biol. Chem. **261**, 8965–8976
6 Boxer, L. M., Miwa, T., Gustafsen, T. A. and Kedes, L. H. (1989) J. Biol. Chem. **264**, 1284–1292
7 Boxer, L. M., Prywes, R., Roeder, R. G. and Kedes, L. H. (1989) Mol. Cell. Biol. **9**, 515–522
8 Kim, J. H., Johansen, F. E., Robertson, N., Catino, J. J., Prywes, R. and Kumar, C. C. (1994) J. Biol. Chem. **269**, 13740–13743
9 Katoh, Y., Loukianov, E., Kopras, E., Zilberman, A. and Periasamy, M. (1994) J. Biol. Chem. **269**, 30538–30545
10 Lees-Miller, J. P., Heeley, D. H. and Smillie, L. B. (1987) Biochem. J. **244**, 705–709
11 Prinjha, R. K., Shapland, C. E., Hsuan, J. J., Totty, N. F., Mason, I. J. and Lawson, D. (1994) Cell Motil. Cytoskeleton **28**, 243–255
12 Kobayashi, R., Kubota, T. and Hidaka, H. (1994) Biochem. Biophys. Res. Commun. **198**, 1275–1280
13 Shanahan, C. M., Weissberg, P. L. and Metcalfe, J. C. (1993) Circ. Res. **73**, 193–204
14 Nishida, W., Kitami, Y., Abe, M. and Kiwada, K. (1991) Biochem. Int. **23**, 663–668
15 Nishida, W. W., Kitami, Y. and Kiwada, K. (1993) Gene **130**, 297–302
16 Thweatt, R., Lumpkin, C. K. and Goldstein, S. (1992) Biochem. Biophys. Res. Commun. **187**, 1–7
17 Takahashi, K. and Nadal-Ginard, B. (1991) J. Biol. Chem. **266**, 13284–13288
18 Kemp, P. R., Grainger, D. J., Shanahan, C. M., Weissberg, P. L. and Metcalfe, J. C. (1991) Biochem. J. **277**, 285–288
19 Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
20 Grainger, D. J., Hesketh, T. R., Metcalfe, J. C. and Weissberg, P. L. (1991) Biochem. J. **277**, 145–151
21 Kemp, P. R., Shachar-Hill, Y., Weissberg, P. L. and Metcalfe, J. C. (1993) FEBS Lett. **336**, 119–123
22 Genetics Computer Group (1991) Program Manual for the GCG Package, Version 7, University of Wisconsin, Madison, WI
23 Osbourn, J. K., Weissberg, P. L. and Shanahan, C. M. (1995) Gene **154**, 249–253
24 Mar, J. H. and Ordahl, C. P. (1990) Mol. Cell. Biol. **10**, 4271–4278
25 Lassar, A. B., Buskin, J. N., Lockshon, D. et al. (1989) Cell **58**, 823–831
26 Weintraub, H., Davis, R., Tapscott, S. et al. (1991) Science **251**, 761–766
27 Ern, H. and Monard, D. (1993) Gene Expr. **3**, 163–174
28 Pfeifer, G.P, Tanguay, R. L., Steigerwald, S. D. and Riggs, A. D. (1990) Genes Dev. **4**, 1277–1287

Received 10 March 1995/1 May 1995; accepted 23 May 1995