Characterization of the human *CREB3L2* gene promoter

IOANNIS PANAGOPOULOS and FREDRIK MERTENS

Department of Clinical Genetics, Lund University Hospital, SE-221 85 Lund, Sweden

Received August 29, 2008; Accepted November 26, 2008

DOI: 10.3892/or_00000264

Abstract. *CREB3L2* encodes a member of the CREB3 family of transcription factors. We characterized its promoter region, showing that it is asymmetrically bidirectional, also driving the expression of a variant of *AKR1D1*. It has a CRE binding site which is conserved among mammalians; removal or alteration of it resulted in reduced promoter activity. When transiently transfecting the HEK293 cell line with constructs with partially deleted promoter regions, 5' deletions beyond 1058-bp upstream of the transcription starting site resulted in successive reduction of the activity. The inclusion of the untranslated part of *CREB3L2* exon 1 strongly inhibited the promoter activity. Forskolin resulted in a decreased reporter activity, whereas phorbol 12-myristate 13-acetate increased the promoter activity irrespective of the status of the CRE binding site. The presence of the CRE site indicates autoregulation of CREB3L2 and/or regulation via other members of the CREB3 family or a variety of bZIP transcription factors.

Introduction

CREB3L2 (BBF2H7) was first identified as the partner gene of *FUS* in a chimera found in two cases of low grade fibromyxoid sarcoma (LGFMS) (1). Screening of a larger number of LGFMS demonstrated that virtually all are characterized by a chimeric *FUS/CREB3L2* gene, and that rare cases may display a variant *FUS/CREB3L1* fusion (2,3). Genomic characterization of *CREB3L2* showed that it spans >120-kbp genomic DNA, is composed of 12 exons and is expressed in most of the 24 human tissues examined (1). The CREB3L2 protein is a member of the CREB3 family of transcription factors. Its B-ZIP domain is highly similar to that in the CREB3L1 (OASIS), CREB3L3 (CREB-H), CREB3L4 (CREB4), CREB3 and *Drosophila* BBF-2 transcription factors with 80, 60, 59, 56 and 71% identity, respectively and the leucine zipper motif of *CREB3L2* is similar to that in CREB3L3 and CREB3L4 (pattern L-X_6-C-X_6-L-X_6-L-X_6-L-X_6-L). Downstream of the B-ZIP domain, and similar to CREB3L1, CREB3L2 contains a hydrophobic region, which was predicted to be an α-helical transmembrane domain and is a typical structural feature for the transcription factors activated by regulated intramembrane proteolysis (1,4). These proteins can dimerize with each other, can bind CRE DNA sequences, can activate transcription through box-B and ATF6 elements, and contain a transcriptional activation domain at the N-terminal part. Although the exact cellular role of the *FUS/CREB3L2* fusion in the pathogenesis of LGFMs is still unclear, one can postulate that all the above mentioned functions may be influenced since the B-ZIP encoding domain of *CREB3L2* is fused to the N-terminal part of *FUS* and comes under the control of the FUS promoter. Thus, FUS/CREB3L2 may act as an abnormal transcription factor that undergoes intramembranous proteolysis, allowing the N-terminal FUS domain with the bZIP domain of *CREB3L2* to enter the nucleus and to activate/deregulate a number of target genes.

In a previous study we assessed various cellular outcomes after transfection of NIH3T3 and HEK-293 cells with constructs containing full-length and truncated (*CREB3L2ΔTM* and *FUS/CREB3L2ΔTM*) versions of *CREB3L2* and *FUS/CREB3L2* (5). The minimum region for transcriptional activation by CREB3L2 was seen in constructs containing the first 120 amino acids (aa). Much stronger transcriptional activation was consistently seen for the FUS/CREB3L2 constructs, in which the first 224 aa of CREB3L2 are replaced by the first 212 aa of FUS, than for the corresponding CREB3L2 constructs. Transcriptional activity was achieved through the box-B element, ATF6 and CRE binding sites, as well as the *GRP78* promoter. Proteins encoded by full-length *CREB3L2* and *FUS/CREB3L2* were localized to reticular structures of the cytoplasm, whereas the corresponding, truncated proteins lacking the transmembrane domain and the carboxy-terminal part of CREB3L2 resided within the nucleus. Recently, CREB3L2 was shown to be induced at the translational level during endoplasmatic reticulum stress, suggesting that it might contribute to a late phase of unfolded protein response signalling (4).

In the present study, we characterized the promoter region of the *CREB3L2* gene.

Materials and methods

Vectors. All PCR amplifications were performed in a 50 μl reaction volume containing 1X AccuPrime *Pfx* reaction mix, 1 unit AccuPrime *Pfx* DNA polymerase (Invitrogen, Carlsbad, CA), 0.3 μM of each of the forward and reverse primers (Table I) and 200 ng template DNA. The PCR was run on a

Correspondence to: Dr Ioannis Panagopoulos, Department of Clinical Genetics, University Hospital, SE-221 85 Lund, Sweden
E-mail: ioannis.panagopoulos@med.lu.se

Key words: CREB3L2 gene, promoter, CRE binding site, bidirectional, *AKR1D1* gene

Table I. Plasmid name, primers used for the amplification of promoter region, size of the amplified products and the corresponding promoter region of *CREB3L2*.

Plasmid	Forward primer	Reverse primer	Size (bp)	Promoter region
p-1297/+396	CGGGGTACCGATAGGACCAGACAAATGCAATA	GGAAGATCTCCGCTCTCCAGCACCTCCAT	1693	-1297/+396
p-1058/+396	CGGGGTACCTGTTCGGGGGAATTCCTTTAG	GGAAGATCTCCGCTCTCCAGCACCTCCAT	1454	-1058/+396
p-831/+396	CGGGGTACCTTGACGCATCCTGTTTTCCTGC	GGAAGATCTCCGCTCTCCAGCACCTCCAT	1227	-831/+396
p-580/+396	CGGGGTACCTGAAGAACAATGACAATTTCCT	GGAAGATCTCCGCTCTCCAGCACCTCCAT	976	-580/+396
p-328/+396	CGGGGTACCGCTATTTCATTGGCTGTACCTT	GGAAGATCTCCGCTCTCCAGCACCTCCAT	724	-328/+396
p-177/+396	CGGGGTACCACAGACACTCCCCGGCTACA	GGAAGATCTCCGCTCTCCAGCACCTCCAT	573	-177/+396
p-1297/+1	CGGGGTACCGATAGGACCAGACAAATGCAATA	GGAAGATCTGCTCTGCTCCAGGACCCAGC	1297	-1297/+1
p-1058/+1	CGGGGTACCTGTTCGGGGGAATTCCTTTAG	GGAAGATCTGCTCTGCTCCAGGACCCAGC	1058	-1058/+1
p-831/+1	CGGGGTACCTTGACGCATCCTGTTTTCCTGC	GGAAGATCTGCTCTGCTCCAGGACCCAGC	831	-831/+1
p-580/+1	CGGGGTACCTGAAGAACAATGACAATTTCCT	GGAAGATCTGCTCTGCTCCAGGACCCAGC	580	-580/+1
p-328/+1	CGGGGTACCGCTATTTCATTGGCTGTACCTT	GGAAGATCTGCTCTGCTCCAGGACCCAGC	328	-328/+1
p-177/+1	CGGGGTACCACAGACACTCCCCGGCTACA	GGAAGATCTGCTCTGCTCCAGGACCCAGC	177	-177/+1

PCT-200 DNA Engine (MJ Research, Waltham, MA). The cycling included an initial denaturation at 95˚C for 2 min, followed by 30 cycles of 15 sec at 95˚C, 30 sec at 58˚C, and 2 min at 68˚C, and a final extension for 5 min at 72˚C. All ligations were performed overnight at 16˚C in a 10 μl reaction volume containing 1X Ligase reaction buffer, 5 units T4 DNA ligase (Invitrogen) and 1:3 vector to insert ratio. When purification was required, the DNA fragments were purified using either the QIAquick gel extraction kit or the QIAquick PCR purification kit (Qiagen, Hilden, Germany). For sequence confirmation, the ABI Prism BigDye terminator v1.1 cycle sequencing kit (Applied Biosystems, Foster City, CA) was used. Dye-incorporated DNA fragments were subsequently purified by centrifugation using DyeEX spin column (Qiagen) according to the manufacturer's instructions and were analyzed on an Applied Biosystems model 3100-Avant DNA sequencing system. All the plasmids were prepared using Qiagen's plasmid miniprep kit.

In order to measure simultaneously promoter activity and transfection efficiency, a report vector, pFhRL, was constructed to contain both *Photinus pyralis* (firefly) and *Renilla reniformis* luciferase genes (6). In the pFLhRL vector, the luciferase genes are placed in opposite directions to avoid read-through, 'leaky' transcription and the expression of *Renilla* luciferase is driven by the SV40 early enhancer/promoter. For the construction of pFhRL, a 1327-bp fragment containing the SV40 early enhancer/promoter and *Renilla* luciferase reporter gene was amplified from the pGL4.73 [*hRluc*/SV40] plasmid (Promega) using the forward primer 5'-GAAAGGTCGACCTGTGGA ATGTGTGTCAGT-3' and the reverse primer 5'-CGCGGA TCCAATTACTGCTCGTTCTTCAGC-3', which contain *Sal*I and *Bam*HI restriction sites, respectively. The fragment was digested with *Sal*I and *Bam*HI restriction endonucleases and subcloned between the corresponding sites of the pGL4.10 [*luc2*] vector, downstream of the firefly luciferase gene.

DNAs containing various lengths of the *CREB3L2* promoter region were PCR amplified using the BAC clone RP11-377B19 (AC009263) as template and various primer sets (Table I) and cloned in the pFLhRL vector between *Kpn*I and *Bgl*II restriction sites, upstream of the firefly luciferese cDNA.

For experiments concerning transcription activation through various DNA-binding sites, pCR3.1 (Invitrogen)-based constructs containing the truncated CREB3L2ΔTM (amino acids 1-372; in this form the transmembrane domain and the C-terminal part have been deleted), and FUS/CREB3L2ΔTM (amino acids 1-366 of the FUS/CREB3L2 protein), and ATF6ΔTM (amino acids 1-377) were used. For the pCR3.1-CREB3L2ΔTM construction, a 1116-bp CREB3L2 cDNA fragment was amplified using the forward CGGGATCCGT ATG GAG GTG CTG GAG AGC G and the reverse GTA GAATTCTTA GCC AGC TAA CTT GCA GGT TCG primers with *Bam*HI and *Eco*RI restriction sites, respectively, and the pCR3.1-*CREB3L2* plasmid as template (5). For the construction of pCR3.1-FUS/CREB3L2ΔTM, a 1097-bp cDNA fragment of *FUS/CREB3L2* was amplified using the FUS forward primer GGCGGATCC ATG GCC TCA AAC GAT TAT ACC C, the above mentioned reverse primer and the pCR3.1-FUS/CREB3L2 plasmid as template (5). Subsequently, the amplified cDNA fragments were digested and inserted

between the *Bam*HI and *Eco*RI restriction sites of the pCR3.1 vector. For the construction of pCR3.1-ATF6ΔTM, the forward primer GTCTGGAAGCTT ATG GGG GAG CCG GCT GGG which contained a *Hin*dIII site and the reverse primer GCGGCTCGAG TCT TCG CTT TGG ACT AGG GAC T with an *Xho*I site were used to amplify a 1131-bp *ATF6* cDNA fragment (amino acids 1-377) from a cDNA of the K562 cell line.

Site-directed mutagenesis was performed using the Gene Tailor site-directed mutagenesis system (Invitrogen) according to the manufacturer's instructions. The reverse primer: ACC GAG GAA CTC CTC CTG CTT GAG CGC GGC CCG CTG TC was used together with 1) AAG CAG GAG GAG TTC CTC GGC ACA GAC ACT CC, 2) AAG CAG GAG GAG TTC CTC GGT GTC GTC ACA GAC ACT CC, 3) AAG CAG GAG GAG TTC CTC GGT GTT GTC ACA GAC ACT CC, and 4) AAG CAG GAG GAG TTC CTC GGT GTT TTC ACA GAC ACT CC to introduce the ΔTGACGT, TGTCGT, TGTTGT and TGTTTT mutations respectively, at the CRE site of the promoter in the p-328/+1 plasmid (Table I).

Cell lines and transfection experiments. The cell line HEK293 (obtained from The Banca Cellule e Colture in GMP, Italy) was used for transfection experiments. The cells were cultured in RPMI-1640 medium, supplemented with fetal bovine serum (10%), 2 mM L-glutamine, 100 IU/ml penicillin, and 100 μg/ml streptomycin. Transfections were performed with PolyFect transfection reagent (Qiagen) according to the manufacturer's recommendations (http://www1.qiagen.com/literature/protocols/pdf/TFP03.pdf). Briefly, $5x10^4$ cells were seeded in 96-well plates, and 24 h later they were transfected with the appropriate plasmid combinations using 1 μl PolyFect transfection reagent. For experiments concerning transcription potential of various parts of the *CREB3L2* promoter (deletion constructs) 500 ng of each vector was used. To study the effects of ATF6ΔTM, CREB3L2ΔTM and FUS/CREB3L2ΔTM, 200 ng of the plasmid p-328/+1 were co-transfected together with 5, 50 and 500 ng of pCR3.1-based expression plasmids which contain the above mentioned transcription factors. To study the effects of unfolded protein response stress inducers, 28 h after transfection, tunicamicin (at concentration 2 μg/ml), thapsigargin (at concentration 500 nM) or brefeldin (at concentration 0.5 μg/ml) were added and the cells were further cultured for 16 h. To study the effects of CRE-related activators, forskolin and phorbol 12-myristate 13-acetate (PMA) were used. After 24 h transfection the culture medium was replaced with a new medium which contained 0.1% fetal bovine serum and 10 μM forskolin, or 100 nM PMA or 10 μM forskolin +100 nM PMA. All the above mentioned chemicals were purchased from Sigma-Aldrich.

Dual luciferase assay. Cells were lysed 48 h after transfection in 40 μl of 1X passive lysis buffer (Promega) and assayed for the firefly and *Renilla* luciferase activities using the Dual Luciferase assay system (Promega), according to the manufacturer's instructions. The results were normalized against the *Renilla* luciferase activities. Measurements were performed with the Veritas 96-microplate luminometer (Turner Biosystems, Sunnyvale, CA) using 10 μl cell lysate and 96-well medium binding Lumitrac 200 plates (Greiner bio-one, Kremsmuenster, Austria). Each construct was measured in eight replicas. The results are presented as the median together with the 25th and 75th percentiles. The Mann-Whitney 2-tailed test was used for the statistical analysis using the statistiXL software (http://www.statistixl.com).

In silico analysis. The UCSC Genome Browser was used (http://genome.ucsc.edu/index.html?org=Human). The sequences of human *CREB3L2* (ENSG00000182158) and the orthologues from 16 mammalians were obtained from ensembl (Ensembl release 49 - March 2008; http://www.ensembl.org/index.html). For genomic alignment the MultAlin software (7) was used (http://bioinfo.genopole-toulouse.prd.fr/multalin/multalin.html). The Patch public 1.0 software was used to identify transcription factor binding sites (http://www.gene-regulation.com/cgi-bin/pub/programs/patch/bin/patch.cgi). In addition, the pre-computed whole genome alignments of the Vista suite of programs and databases were used for comparative genomic sequences (http://genome.lbl.gov/vista/index.shtml): Vista browser for the pairwise alignment of the human and mouse sequence, and the Whole Genome rVista in order to evaluate conserved transcription factor binding sites (TFBS) in the upstream region of *CREB3L2*. The repeatmasker was used to identify sequence repeats (http://www.repeatmasker.org/). The sequence together with the first exon of *CREB3L2* was also screened for CpG islands using the CpG Island searcher (http://cpgislands.usc.edu/) (8). Default parameters for the identification of CpG islands included using the lower-limit values 500 bp for length, 55% for GC content, and 0.65 for Observed CpG/Expected CpG (8).

Expression analysis. Expression analyses (RT-PCR) were performed using testis, liver and colon ready cDNA obtained from human MTC panels I and II (Clontech, Mountain View, CA), the cell lines TERA-2 (HTB-106) and NCCIT (CRL-2073) obtained from LGC Promochem (http://www.lgcpromochem-atcc.com/) and an ovarian dysgerminoma from a 13-year-old girl. Total RNA from the two mentioned cells lines and the dysgerminoma was extracted using the TRIzol reagent (Invitrogen) and cDNA synthesis was done as described (1). The primer combination AKR1AK-1F (AAT AGA GGG CTT GGG TCC GGA AGT) and AKR1D1AK-396R (TTG GGT CCG GAA GTG ATG GAC TG) was used for the amplification. The 50 μl reaction volume contained 20 mM Tris-HCl pH 8.4 (at 25˚C), 50 mM KCl, 1.25 mM $MgCl_2$, 0.2 mM of each dNTP, 1 unit Platinum*Taq* DNA polymerase (Invitrogen), 0.5 μM of each of the forward AKR1AK-1F and reverse AKR1D1AK-396R primers, and 3 μl of the ready cDNA or 2 μl cDNA from the cell line. PCR was run using a PCT-200 DNA engine (MJ Research) and the cycling was: an initial cycle at 94˚C for 2 min, 35 cycles of 1 min at 94˚C, 1 min at 60˚C, and 1 min at 72˚C, followed by a final extension for 10 min at 72˚C. PCR products (15 μl) were analyzed by electrophoresis through 1.5% agarose gels, stained with ethidium bromide, and photographed. The amplified DNA fragments were further purified using QIAquick gel extraction kit (Qiagen), directly sequenced using the ABI Prism BigDye terminator v1.1 cycle sequencing kit (Applied Biosystems) and analyzed on an Applied Biosystems model 3100-Avant DNA sequencing system. The

Table II. Summary of all conserved tanscription factor binding sites located in a 2000-bp upstream region of *CREB3L2* (locus link 64764) identified with the engine Whole Genome rVista.

Number of hits	Transcription factor
1	AREB6, ATF3, BARBIE, CAAT, CDPCR3HD, CREBATF, CREBP1CJUN, DR3, E4F1, ER, FOXM1, FOXO3A, GATA3, HAND1E47, HFH3, HNF4ALPHA, IK3, IRF1, LEF1TCF1, MMEF2, NKX3A, NKX61, NRF2, PAX6, PITX2, POU1F1, PPARG, RBPJK, RFX1, SOX9, TEF1, XBP1
2	AP1, ATF, ATF1, ATF4, BACH2, CEBPDELTA, CEBPGAMMA, CETS1P54, CREBP1, E2F1DP1, E2F1DP2, E2F4DP2, ELK1, FOXJ2, FREAC7, GATA4, HELIOSA, HNF3B, HOXA4, KROX, LDSPOLYA, MAZ, MAZR, MINI19, MYCMAX, NFAT, PBX, POU6F1, SZF11, TBX5, TFIII, XPF1
3	CHOP, HLF, NFY, SF1, SOX5, TFIIA
4	ETS, IRF7, STAT, VMAF
5	MRF2, MTATA
6	E2F1, FOX, IRF
7	CDX2, CP2, LEF1
8	CDC5
9	CREB, EFC, SRF
10	MEF2
11	CDX, CMAF, EGR, PIT1
12	HNF3ALPHA
13	XVENT1
15	CREL
16	FOXD3, PAX3
17	HNF3, NFKB, TST1
19	CACD, TATA
21	GC
24	OCT4, PAX
27	OCT1
31	E2F, PAX4
33	LYF1
36	EVI1
37	VDR
46	DR4, HNF1, HNF4
59	PPAR
61	AP2
82	DR1
83	SP1

BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat? command=start) and BLAST (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) programs were used for computer analyses of sequence data.

Results

In silico and expression analyses. A 2000-bp DNA fragment upstream from the first exon of *CREB3L2* was chosen arbitrarily for promoter analysis. The DNA sequence was first screened for interspersed repeats and low complexity DNA sequences using the RepeatMasker program. The analysis identified a 305-bp ALU repeat, a 259-bp LINE2 fragment, and a 209-bp MER2 DNA element positioned 1296-1600 bp, 686-975 bp and 1791-2000 bp upstream of *CREB3L2* exon 1, respectively (Fig. 1). In addition, the sequence together with exon 1 of *CREB3L2* has one 974-bp CpG-island with 60% CG content (from position-477 through entire exon 1) and ObservedCpG/ExpectedCpG ratio 0.785. Using the Patch public 1.0 software and the Whole Genome rVista 140 conserved transcription factor binding site (TFBS) were identified (Table II). Furthermore, the Whole Genome rVista program showed that CRE binding site (BS), GTGACGTCAC, was highly conserved (Fig. 1). Two other conserved BS were also identified with TFBS conserved track in UCSC genome browser: a TGIF site within exon 1 of *CREB3L2* and an NKX6-1 site, 139-bp upstream of the CRE site (Fig. 1).

Alignment of the human (ENSG00000182158) and the orthologues sequences from 16 mammalians revealed high conservation of the CRE BS (Table III). Interestingly, 18 nucleotides (20 for opposum) downstream of the CRE BS an ACATTCC sequence motif was found to be conserved in all mammalians. The ACATTCCA sequence was reported to be a binding site of miR-206, miR-1 and miR-122a (9). No TATA-box was found but at position -97 to -90 a palindromic CCAATTGG was observed.

According to the UCSC genome browser, 283-bp upstream of the +1 transcription site and in opposite direction to *CREB3L2* a cDNA, AK058142, has been mapped. The sequence is specific for testis, is supported by four other spliced EST sequences (BU561106, AI198697, AI656277 and AI652102) and seems to be an alternative variant of *AKR1D1* which codes for aldo-keto reductase family 1, member D1 (δ4-3-ketosteroid-5-ß-reductase). RT-PCR with the AKR1AK-1F forward primer and the AKR1D1AK-396R reverse primer amplified three fragments from the testis cDNA but none from liver, colon, TERA-2, NCCIT or the dysgerminoma (Fig. 2). Direct sequencing of the amplified fragments showed that they were alternative transcript variants of the *AKR1D1* gene and contained parts of *AK058142, AKR1D1* (NM_005989) and an additional sequence presumably represented by not previously defined exons specifically expressed in testis (Fig. 2). From the long 1092-bp transcript (GenBank accession no. FM160741) the entire AKR1D1 (NP_005980) would be translated. The 930 bp (GenBank accession no. FM160742) would result in an AKR1D1 protein with an alternative N-terminal; the first 32 amino acids (MDLSAASHRIPLSDGNSIPIIGLGTYSEP KST) would be replaced by 42 amino acids (MLGSLAWD AISLTPNAWLNSSAALGLSSVNHSIASLYAGVWP), while the 566-bp fragment (GenBank accession no. FM160743)

Table III. Partial alignment in the promoter region of 17 mammalians.[a]

Species	Ensembl Gene ID	Position	Sequence					
Human	ENST00000330387	-575 to -518	GGAGGAGTTC	CTCGGTGACG	TCACA--GAC	ACTCCCCGGC	TACATTCCAA	GGCGTGGGCG
Mouse lemour	ENSMICG00000003960	-583 to -526	GGAGGAGTTC	CCGGGTGACG	TCAAA--GGC	ACTGCCCGCC	CACATTCCAA	AGCCTGGGCG
Squirrel	ENSSTOT00000009816	-571 to -514	GGAGGAGTTC	CCCGGTGACG	TCAAG--GGC	GCTCCCCGCC	CACATTCCAA	GGCATGGGCG
Cow	ENSBTAG00000015802	-586 to -529	AGAGTAGTTC	CCCGGTGACG	TCAAA--GGC	GCTGCCTGCC	CACATTCCAA	GGCATGGGCG
Horse	ENSECAG00000007609	-590 to -533	GGAGGAGTTC	TCCGGTGACG	TAAAC--GGC	ACGCCCTGCC	CACATTCCAA	AGCCTGGGCG
Cat	ENSFCAG00000000648	-591 to -534	GAAGGAGTTC	CCCGGTGACG	TCAAA--GGC	GCTCCCCGCC	CACATTCCAA	AGTCTGGGCA
Pika	ENSOPRG00000004323	-604 to -547	AGAGGAGTTC	CCCGCTGACG	TCAAA--GGC	GCACCCGGCC	CACATTCCAA	AGACTAGGCC
Armadillo	ENSDNOG00000004102	-560 to -503	GGAGGTGTTC	CACGGTGACG	TCAAA--GAA	GCGCCCCTCC	CACATTCCAA	GGCCTCGGCG
Tree shrew	ENSTBET00000004884	-569 to -512	GGAGGAGTTT	CTCAGTGACG	TCAGA--GGC	ACTCCCCGCC	CACATTCCAA	AGGTTGGGCG
Dog	ENSCAFG00000013481	-563 to -506	GAGGGAGTTC	CCCGGTGACG	TCAGA--GGC	GCTCCCCGCC	CACATTCCAA	AGCCTGGGCT
Bushbaby	ENSOGAT00000010848	-540 to -483	GGAGGAGTTC	CCCGGTGACG	TCAAA--GTC	ACGCCTCGCC	TACATTCCAA	AGCCGGGGCG
Mouse	ENSMUSG00000038648	-559 to -502	GGAGGCGTTC	CCCAGTGACG	TAAAA--GAC	CCTCTCCGCC	CACATTCCGA	GGACTAGGCG
Rabbit	ENSOCUT00000007313	-505 to -448	GGAGGAGTTC	CCTGGTGACG	TCCAG--GGC	GCTCCCCGCC	CACATTCCAA	AGACTAGGCG
Guinea pig	ENSCPOG00000008872	-514 to -457	GGAGGAGTTC	CCCGGTGACG	TCAAA--GGT	GCTCCCAGCC	CACATTCCAA	AGCCTGGGCG
Shrew	ENSSART00000000499	-564 to -508	GGAGGCGGTC	CTCGGTGACG	TCAAA--G-C	GCTCCCCGCC	TACATTCCAC	AGACTGGGCG
Opposum	ENSMODG00000013876	-650 to -591	GAAGGAGTTA	CCCGGTGACG	TCAAACCGGG	ACTCCCCTCC	CACATTCCTA	GGTCGGGGCG
Elephant	ENSLAFG00000008469	-480 to -423	GGAGGAGTTC	GCTGGTGACG	TCAAA--GGC	GCGCCCCGCC	CACATTCCAC	AGCCGGGGCG
Consensus			GgAGGaGTTc	cccgGTGACG	TcAaa Ggc	gCtCCCcgCC	cACATTCCaa	aGcctgGGCg

[a]The sequences are extracted from Ensembl release 49-Mar 2008. Position is calculated upstream from the first ATG codon. The consensus GTGACGT-18N-ACATTCC is found in all species but opposum and shrew.

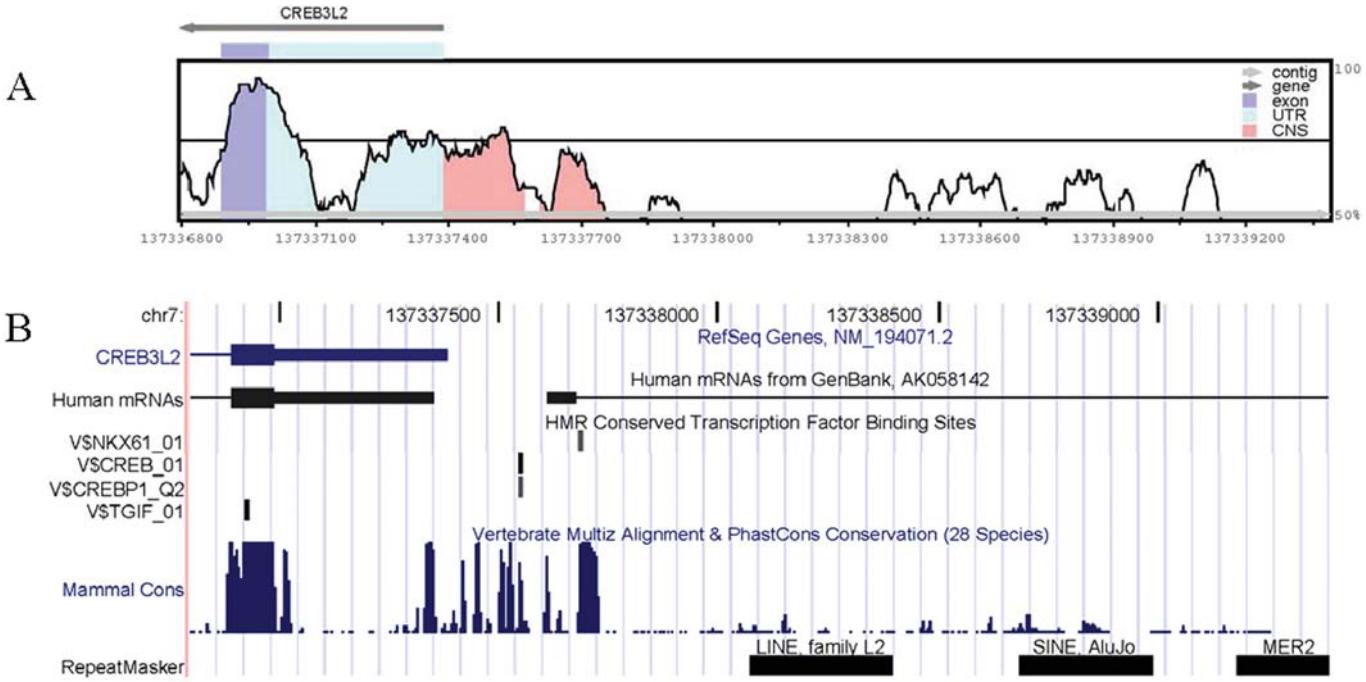


Figure 1. The promoter region of *CREB3L2*. (A) Alignment of mouse chr6 (+) 37371228-37373645 with human chr7:137336800-137339386 (X-axis). The pairwise alignment of Human March 2006 genome with the Mouse February 2006 genome was performed with Vista browser. (B) Diagram, obtained from the UCSC genome browser on Human March 2006 assemby, showing *CREB3L2, AK058142*, conserved transcription factor binding sites, evolutionary conserved DNA regions and interspersed DNA repeats. CNS indicates conserved non-coding sequence.

would correspond to a shorter form of AKR1D1 lacking the first 153 amino acids.

Deletion analysis of the CREB3L2 promoter. To obtain experimental evidence and identify regions that are important for promoter activity of the 5'-flanking sequences, a series of partial deletion promoter constructs was used to transiently transfect HEK293 cell line. The highest activity was obtained with the constructs p-1297/+1 and p-1058/+1 (Fig. 3). Further 5' deletions resulted in successive reduction of the promoter activity. Compared to the p-1297/+1 construct activity reduction in the activity by 10, 17, 30 and 60% was observed for p-831/+1, p-580/+1, p-328/+1 and p-177/+1, respectively. The inclusion of the untranslated part of *CREB3L2* exon 1 strongly inhibited the promoter activity (Fig. 3), suggesting the presence of negative regulatory elements.

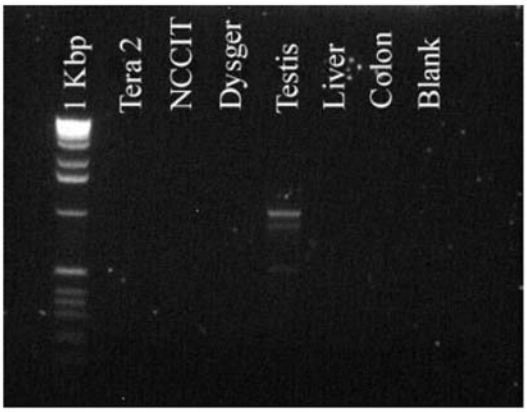
Analysis of CRE binding site. To test the functional importance of the CRE BS, GTGACGTCAC, directed mutagenesis was used to selectively disrupt the core of the site. Thus, the CRE BS of the promoter in the p-328/+1 plasmid was completely removed or 1, 2 or 3 nucleotides were altered generating ΔTGACGT, TGTCGT, TGTTGT and TGTTTT mutations, respectively. Compared to the p-328/+1 construct, a progressive reduction of the promoter activity by 16, 27 and 35% was obtained with substitution of one, two or three nucleotides, respectively (Fig. 4A). Complete removal of the binding site resulted in the same reduction (33%) as the replacement of three nucleotides (Fig. 4A).

To study the effects of the ATF6ΔTM, CREB3L2ΔTM and FUS/CREB3L2ΔTM transcription factors on the *CREB3L2* promoter, 200 ng of the plasmid p-328/+1 was co-transfected together with 5, 50 and 500 ng of pCR3.1-based expression plasmids which contain the above mentioned transcription factors. The analysis showed that there was an increase in the promoter activity with increased amount of the co-transfected ATF6ΔTM and CREB3L2ΔTM transcription factors (Fig. 4B), while FUS/CREB3L2ΔTM had virtually the same effect as the empty pCR3.1 vector on the promoter activity. Deletion of the CRE BS resulted in a 30% reduction of the promoter activity induced by CREB3L2ΔTM but only in 10% reduction of the reporter activity induced by ATF6ΔTM (Fig. 4C).

Effects of ER stress inducers brefeldin, thapsigargin and tunicamycin and the CRE-related activators forskolin and phorbol 12-myristate 13-acetate (PMA) on the promoter activity of CREB3L2. In cells transfected with p-328/+1 plasmid, tunicamicin increased the reporter activity by 11% (p=0.05), thapsigargin by 14% (p=0.05) and brefeldin by 31% (p=0.0001) compared to the control (Fig. 5A). However, in cells transfected with ΔTGACGT or TGTTGT there was not any statistically significant difference in the reporter activity between control and the treatment with tunicamicin, thapsigargin or brefeldin (Fig. 5A).

Culture of HEK293 cells with 0.1% serum increased the reporter activity for all constructs compared to culture with 10% serum. Thus, there was a 30% increase in the activity of p-328/+1 (p=0.0003) and ΔTGACGT (p=0.05) and a 12% increase in the activity of TGTTGT (p=0.028) (Fig. 5B). Treatment of cells with forskolin resulted in a decreased reporter activity, whereas treatment with PMA increased the promoter activity irrespectively of the presence, mutation or

A



B

```
   1 AATAGAGGGC TTGGGTCCGG AAGTGATGGA CTGGAGGCTT GACTAATGTG AAAGGGGCGT
  61 CTGAAAGAGA AGCCAGCTGC CATGCCATGA ACTGCAATGT GAAAAGGCCA ATGTAGCAAG
 121 GATATGTCCA GCCAACAGCC AGTGAGGATC TGAGGCCCTC AGACCAGCAG CCCATGAGCA
 181 ACTGAATTTT GCTAATAACC ATGTAAGTGA ACACAGAAGT GGATCCTTCC CCAGGCGAGC
 241 CTTCATAAGT CCACAGACCC AGTCAAGACT TTGACTGCAG CCTTGTGAGA GCCCCTAAGC
 301 CAGAGGACCC AGCTAAGCCA CGTTCAACTT CCTGACCTGT CAGATCTTGG CGGTTCAATC
 361 TGGAAACTTG GTGTGCTGCG TGTGAATACA GATGCTGGGA AGCTTAGCCT GGGATGCCAT
 421 CTCTCTGACC CCTAATGCTT GGCTGAACTC TTCTGCTGCA CTCGGTCTTT CCTCTGTAAA
 481 TCACAGCATT GCATCACTAT ACGCTGGAGT CTGGCCCTAG GACACCTTTC TAAAAAGACT
 541 CCCTGTGGTG TTCAGAATCA CTCCTACAGT CAGGTTCTCC ACAATGGATC TCAGTGCTGC
 601 AAGTCACCGC ATACCTCTAA GTGATGGAAA CAGCATTCCC ATCATCGGAC TTGGTACCTA
 661 CTCAGAACCT AAATCGACCC CTAAGGGAGC CTGTGCAACA TCGGTGAAGG TTGCTATTGA
 721 CACAGGGTAC CGACATATTG ATGGGGCCTA CATCTACCAA AATGAACACG AAGTTGGGGA
 781 GGCCATCAGG GAGAAGATAG CAGAAGGAAA GGTGCGGAGG GAAGATATCT TCTACTGTGG
 841 AAAGCTATGG GCTACAAATC ATGTCCCAGA GATGGTCCGC CCAACCCTGG AGAGGACACT
 901 CAGGGTCCTC CAGCTAGATT ATGTGGATCT TTACATCATT GAAGTACCCA TGGCCTTTAA
 961 GCCAGGAGAT GAAATATACC CTAGAGATGA GAATGGCAAA TGGTTATATC ACAAGTCAAA
1021 TCTGTGTGCC ACTTGGGAGG CGATGGAAGC TTGCAAAGAC GCTGGCTTGG TGAAATCCCT
1081 GGGAGTGTCC AA
```

C

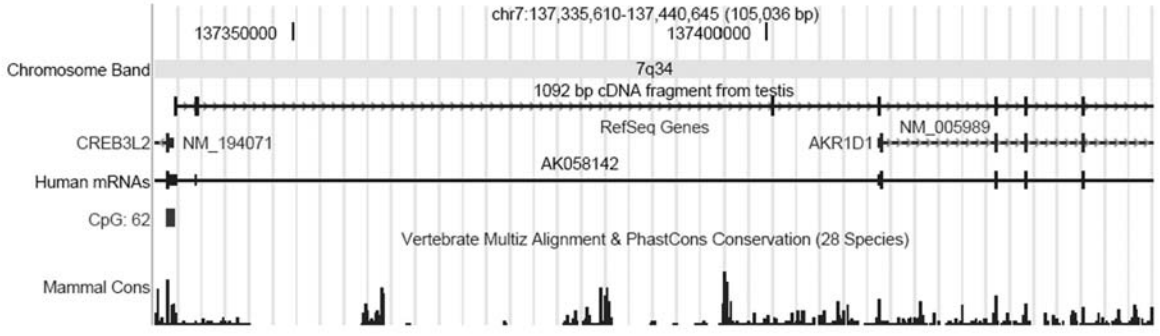


Figure 2. The transcript variants of *AKR1D1* found in testis. (A) Gel electrophoresis showing that cDNA fragments are amplified only from testis. (B) Sequence of the long 1092-bp fragment. The sequences from 415 to 670 and 415 to 1030 are absent in the middle and short amplified fragments, respectively. The three ATG starting codons are in box. (C) Diagram showing the position of the long 1092-bp cDNA fragment, the AK058142 mRNA and the reference sequences of *CREB3L2* and *AKR1D1*. GenBank accession nos. FM160741, FM160742, FM160743.

absence of the CRE binding site (Fig. 5B). In the cells treated simultaneously with both forskolin and PMA the promoter activity was similar to that observed in cells cultured with 0.1% serum (Fig. 5B).

Discussion

In the present study, the promoter region of *CREB3L2*, a gene widely expressed in many tissues, was characterized. The

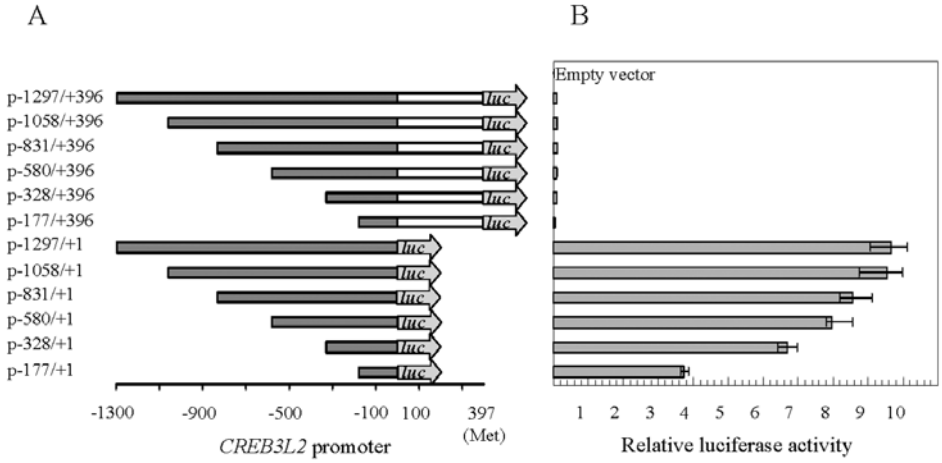


Figure 3. Constructs and functional analysis of the *CREB3L2* promoter. (A) Schematic representation of the various deletion fragments cloned upstream of the luciferase coding sequence (B). Transcriptional activities measured as relative luciferase activity. HEK293 cells were transiently transfected with the reporter plasmids and 48 h after transfection luciferase activity was quantified and normalized against *Renilla* luciferase activities. Each construct was measured in eight replicas and the results are presented as the median together with 25th and 75th percentiles.

sequence CCAATTGG found a few nucleotides upstream of the +1 transcription site of *CREB3L2* suggested that the CCAAT-box could be used bidirectionally. Information obtained from the database and supported by our RT-PCR analysis (Fig. 2) indicated that the promoter might be asymmetrically bidirectional, exhibiting strong directionality towards the *CREB3L2* transcript initiation, with expression of *AKR1D1* alternative transcripts only in testis. *AKR1D1* codes for the enzyme 5ß-reductase which catalyzes the reduction of the 4-ene of 3-ketosteroids, converting them into 5ß-dihydro-3-ketosteroids (10) and is highly expressed in liver and weakly in testis and colon (10).

Bidirectional promoters have been reported for >10% of the genes in the human genome, with the majority of them showing activity in both directions. About 10% seem to function only in one direction suggesting that these promoters drive the expression of one gene while inhibiting expression in the other direction (11). The impact, if any, of expression of *AKR1D1* on *CREB3L2* is unknown. In general, bidirectional promoters lack TATA boxes, are both GC-rich and enriched in CpG islands, and have an overrepresented set of motifs, including GABPA, MYC, E2F1, E2F4, NRF-1, CCAAT, YY1, and ACTACAnnTCC (12). GABPA, E2F1 and CCAAT

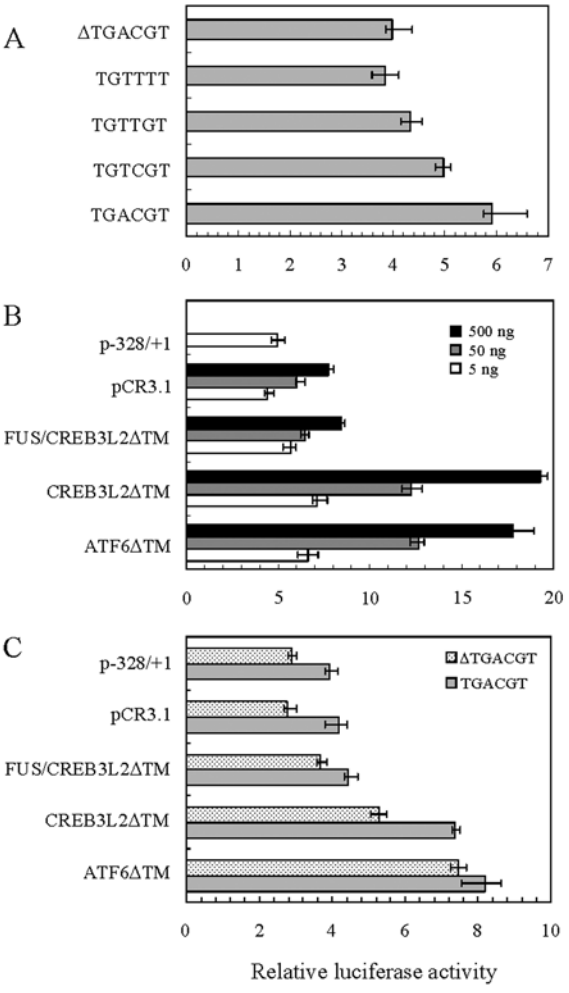


Figure 4. Analysis of the CRE binding site. (A) Site-directed mutagenesis was performed on the p-328/+1 plasmid to completely remove the CRE binding site of the promoter or to alter 1, 2 and 3 nucleotides generating ΔTGACGT, TGTCGT, TGTTGT and TGTTTT mutations, respectively. Compared to the p-328/+1 construct, a progressive reduction of the promoter activity by 16, 27 and 35% was obtained with substitution of one, two or three nucleotides, respectively. Complete removal of the binding site resulted in the same reduction (33%) as the replacement of three nucleotides. (B) Effects of the ATF6ΔTM, CREB3L2ΔTM and FUS/CREB3L2ΔTM transcription factors on the *CREB3L2* promoter. Plasmid p-328/+1 (200 ng) were co-transfected with 5, 50 and 500 ng of pCR3.1-based expression plasmids which contain the above mentioned transcription factors. There was an increase in the reporter activity with increased amount of the co-transfected ATF6ΔTM and CREB3L2ΔTM transcription factors, while FUS/CREB3L2ΔTM had virtually the same effect as the empty pCR3.1 vector on the promoter activity. (C) Effects of the ATF6ΔTM, CREB3L2ΔTM and FUS/CREB3L2ΔTM transcription factors on the p-328/+1 construct without CRE site (ΔTGACGT). Plasmid p-328/+1 (200 ng) were co-transfected with 50 of pCR3.1-based expression plasmids. Deletion of the CRE site resulted in a reduction of the promoter activity induced by CREB3L2ΔTM but did not abolish the reporter activity induced by ATF6ΔTM. HEK293 cells were transiently transfected with the reporter plasmids and 48 h after transfection luciferase activity was quantified and normalized against *Renilla* luciferase activities. Each construct was measured in eight replicas and the results are presented as the median together with 25th and 75th percentiles.

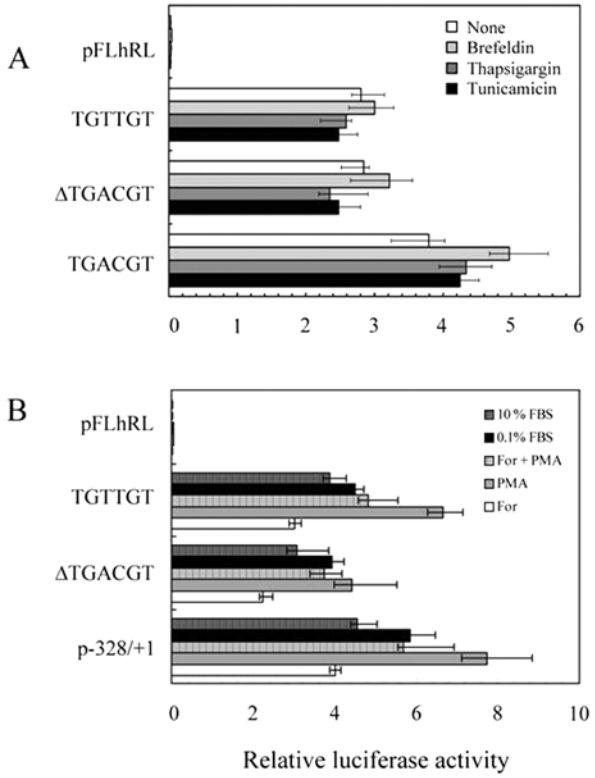


Figure 5. (A) Effects of ER stress inducers brefeldin, thapsigargin and tunicamycin on the promoter activity of CREB3L2. Apart from the the p-328/+1 plasmid with the wild-type CRE (TGACGT) binding site, plasmids with ΔTGACGT and TGTTGT were studied. In cells transfected with p-328/+1 plasmid and compared to control, tunicamicin increased the reporter activity 11% (p=0.05), thapsigargin by 14% (p=0.05) and brefeldin by 31% (p=0.0001). In cells transfected with ΔTGACGT and TGTTGT there was not any statistical significant difference between the reporter activity of the control and cells treated with tunicamicin, thapsigargin or brefeldin. B. Effects of CRE related activators forskolin (For) and phorbol 12-myristate 13-acetate (PMA). HEK293 cells transfected with p-328/+1, ΔTGACGT or TGTTGT were cultured in a medium with 0.1% fetal bovine serum with forskolin, PMA, forskolin + PMA or without them. Culture of HEK293 cells with 0.1% serum (FBS) increased the reporter activity for all constructs compared to culture with 10% serum. Treatment of cells with forskolin resulted in a decreased reporter activity, whereas treatment PMA increased the promoter activity irrespective of the presence, mutation or absence of the CRE binding site. In the cells treated simultaneously with both forskolin and PMA the promoter activity was similar to that observed in cells cultured with 0.1% serum.

motifs are present in the region between +1 of *CREB3L2* and +1 of the *AK058142* trancsript. Moreover, the 10-mer palindromic, GTGACGTCAC, CRE BS found in this region can regulate the transcription in both directions. Interestingly, of the mammalians presented in Table III only humans have the 10-mer palindromic sequence. The presence of the CRE BS suggests that *CREB3L2* is a member of the 'CREB regulon' (13).

The CRE BS is known as a cAMP responsive element and is bound by a variety of bZIP transcription factors including CREB, ATF1, ATF6, ATF2/JUN heterodimers and all members of the CREB3 family of the transcription factors (4,14,15). Thus, autoregulation of CREB3L2 might be suggested as well as regulation via the widely expressed CREB3L1, the liver-specific CREB3L3 and testis- and prostate-specific CREB3L4. Since FUS/CREB3L2ΔTM had virtually no effect on the promoter of *CREB3L2* a function of the chimera could be a disruption of CREB3L2 autoregulation. To study the effects of CRE-related activators, forskolin, which activates adenylate cyclase and results in an increase of intracellular cyclic AMP, and PMA which activates protein kinase C, were used.

The negative effect of forskolin, irrespective of the presence or absence of TGACGT, suggested that other putative binding sites for the CRE family might be present and that cAMP may serve as a negative regulator of the *CREB3L2* transcript. Negative effects of forskolin have been reported for other gene promoters such as *GYS1, CD38*, and the insulin promoter (16-18). Similarly to the *CREB3L2* promoter, Ding *et al* (18) showed that mutagenesis of the CRE-like element in the insulin promoter had no effect on the forskolin-induced suppression. PMA induced promoter activity of *CREB3L2* irrespective of the presence or absence of the CRE site. Indeed, PMA function might involve Egr1, AP1 and Sp1 transcription factor sites found in the promoter of *CREB3L2* (19). Interestingly, forskolin and PMA have opposite effects on the *CREB3L2* promoter. ER stress was shown to increase the amount of total CREB3L2 protein by more than ten-fold and the mRNA level by two-fold (4). To study the effects of ER stress on the promoter activity of CREB3L2, tunicamicin (an inhibitor of N-linked glycosylation), thapsigargin (which blocks the ER calcium ATPase pump, leading to depletion of ER calcium stores) and brefeldin (which specifically blocks protein transport from ER to the Golgi apparatus) were used. All three ER stress inducers increased promoter activity but not to the same magnitude as reported for the protein level. Our data indirectly support the conclusion drawn by Kondo *et al* that the CREB3L2 protein is up-regulated at the translational level during ER stress (4).

In conclusion, our data show that the expression driven by the promoter of CREB3L2 is, at least in part, regulated by CRE binding transcription factors. Consequently, this regulation would be disrupted in low grade fibromyxoid sarcoma with the FUS/CREB3L2 fusion, which in turn might be significant for tumor development.

Acknowledgements

We thank Margareth Isaksson for her technical assistance. This work was supported by The Swedish Childhood Cancer Foundation, The Swedish Research Council and Gunnar Nilsson Cancer Foundation.

References

1. Storlazzi CT, Mertens F, Nascimento A, Isaksson M, Wejde J, Brosjö O, Mandahl N and Panagopoulos I: Fusion of the FUS and BBF2H7 genes in low grade fibromyxoid sarcoma. Hum Mol Genet 12: 2349-2358, 2003.
2. Panagopoulos I, Storlazzi CT, Fletcher CD, Fletcher JA, Nascimento A, Domanski HA, Wejde J, Brosjö O, Rydholm A, Isaksson M, Mandahl N and Mertens F: The chimeric FUS/CREB3L2 gene is specific for low-grade fibromyxoid sarcoma. Genes Chromosomes Cancer 40: 218-228, 2004.
3. Mertens F, Fletcher CD, Antonescu CR, Coindre JM, Colecchia M, Domanski HA, Downs-Kelly E, Fisher C, Goldblum JR, Guillou L, Reid R, Rosai J, Sciot R, Mandahl N and Panagopoulos I: Clinicopathologic and molecular genetic characterization of low-grade fibromyxoid sarcoma, and cloning of a novel FUS/CREB3L1 fusion gene. Lab Invest 85: 408-415, 2005.

4. Kondo S, Saito A, Hino S, Murakami T, Ogata M, Kanemoto S, Nara S, Yamashita A, Yoshinaga K, Hara H and Imaizumi K: BBF2H7, a novel transmembrane bZIP transcription factor, is a new type of endoplasmic reticulum stress transducer. Mol Cell Biol 27: 1716-1729, 2007.
5. Panagopoulos I, Möller E, Dahlén A, Isaksson M, Mandahl N, Vlamis-Gardikas A and Mertens F: Characterization of the native CREB3L2 transcription factor and the FUS/CREB3L2 chimera. Genes Chromosomes Cancer 46: 181-191, 2007.
6. Möller E, Isaksson M, Mandahl N, Mertens F and Panagopoulos I: Comparison of the proximal promoter regions of the PAX3 and PAX7 genes. Cancer Genet Cytogenet 178: 114-119, 2007.
7. Corpet F: Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16: 10881-10890, 1988.
8. Takai D and Jones PA: The CpG island searcher: a new WWW resource. In Silico Biol 3: 235-240, 2003.
9. Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES and Kellis M: Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. Nature 434: 338-345, 2005.
10. Charbonneau A and The VL: Genomic organization of a human 5beta-reductase and its pseudogene and substrate selectivity of the expressed enzyme. Biochim Biophys Acta 1517: 228-235, 2001.
11. Trinklein ND, Aldred SF, Hartman SJ, Schroeder DI, Otillar RP and Myers RM: An abundance of bidirectional promoters in the human genome. Genome Res 14: 62-66, 2004.
12. Lin JM, Collins PJ, Trinklein ND, Fu Y, Xi H, Myers RM and Weng Z: Transcription factor binding and modified histones in human bidirectional promoters. Genome Res 17: 818-827, 2007.
13. Impey S, McCorkle SR, Cha-Molstad H, Dwyer JM, Yochum GS, Boss JM, McWeeney S, Dunn JJ, Mandel G and Goodman RH: Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. Cell 119: 1041-1054, 2004.
14. Hai T and Hartman MG: The molecular biology and nomenclature of the activating transcription factor/cAMP responsive element binding family of transcription factors: activating transcription factor proteins and homeostasis. Gene 273: 1-11, 2001.
15. Wang Y, Shen J, Arenzana N, Tirasophon W, Kaufman RJ and Prywes R: Activation of ATF6 and an ATF6 DNA binding site by the endoplasmic reticulum stress response. J Biol Chem 275: 27013-27020, 2000.
16. Fredriksson J, Ridderstrale M, Groop L and Orho-Melander M: Characterization of the human skeletal muscle glycogen synthase gene (GYS1) promoter. Eur J Clin Invest 34: 113-121, 2004.
17. Sun L, Iqbal J, Zaidi S, Zhu LL, Zhang X, Peng Y, Moonga BS and Zaidi M: Structure and functional regulation of the CD38 promoter. Biochem Biophys Res Commun 341: 804-809, 2006.
18. Ding WQ, Dong M, Ninova D, Holicky EL, Stegall MD and Miller LJ: Forskolin suppresses insulin gene transcription in islet beta-cells through a protein kinase A-independent pathway. Cell Signal 15: 27-35, 2003.
19. Eyries M, Agrapart M, Alonso A and Soubrier F: Phorbol ester induction of angiotensin-converting enzyme transcription is mediated by Egr-1 and AP-1 in human endothelial cells via ERK1/2 pathway. Circ Res 91: 899-906, 2002.