UC Riverside
UC Riverside Previously Published Works

Title

Core promoter-specific gene regulation: TATA box selectivity and Initiator-dependent bi-directionality of serum response factor-activated transcription

Authors

Xu, Muyu
Gonzalez-Hurtado, Elsie
Martinez, Ernest

Core promoter-specific gene regulation: TATA box selectivity and Initiator-dependent bi-directionality of serum response factor-activated transcription

Muyu Xu [a,1], Elsie Gonzalez-Hurtado [a,b,2], Ernest Martinez [a,b,*]

[a] Department of Biochemistry, University of California, Riverside, CA 92521, USA
[b] MARC U-STAR Program, University of California, Riverside, CA 92521, USA

ABSTRACT

Gene-specific activation by enhancers involves their communication with the basal RNA polymerase II transcription machinery at the core promoter. Core promoters are diverse and may contain a variety of sequence elements such as the TATA box, the Initiator (INR), and the downstream promoter element (DPE) recognized, respectively, by the TATA-binding protein (TBP) and TBP-associated factors of the TFIID complex. Core promoter elements contribute to the gene selectivity of enhancers, and INR/DPE-specific enhancers and activators have been identified. Here, we identify a TATA box-selective activating sequence upstream of the human β-actin (ACTB) gene that mediates serum response factor (SRF)-induced transcription from TATA-dependent but not INR-dependent promoters and requires the TATA-binding/bending activity of TBP, which is otherwise dispensable for transcription from a TATA-less promoter. The SRF-dependent ACTB sequence is stereospecific on TATA promoters but activates in an orientation-independent manner a composite TATA/INR-containing promoter. More generally, we show that SRF-regulated genes of the actin/cytoskeleton/contractile family tend to have a TATA box. These results suggest distinct TATA-dependent and INR-dependent mechanisms of TFIID-mediated transcription in mammalian cells that are compatible with only certain stereospecific combinations of activators, and that a TBP-TATA binding mechanism is important for SRF activation of the actin/cytoskeleton-related gene family.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Proper development and homeostasis of eukaryotic organisms rely on the accurate spatio-temporal regulation of specific gene transcription in response to diverse physiological and environmental signals. Specific gene transcription by RNA polymerase II (RNAPII) is controlled by the combinatorial assortment of a variety of regulatory DNA sequences located at promoter-proximal and more distal (i.e., enhancer) regions. These regulatory DNA elements are bound by sequence-specific regulators (activators or repressors) that in turn recruit a variety of coregulators (coactivators or corepressors) to influence transcription initiation by the RNAPII basal transcription machinery at the core promoter. Hence, the core promoter and the basal transcription machinery are the ultimate targets of signal-dependent regulator/coregulator complexes [1,2].

The core promoter is the minimal DNA sequence surrounding the transcription start site (TSS) that is sufficient to direct a low (i.e., basal) level of accurate transcription initiation by the basal RNAPII transcription machinery [1]. The canonical basal transcription machinery for RNAPII is composed of six basal (or "general") transcription initiation factors: TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH, which interact with core promoter DNA and RNAPII to form the preinitiation complex (PIC). Core promoters are structurally diverse and may contain a number of DNA sequence elements that are bound by the basal transcription factors, and primarily by TFIID and TFIIB, during the assembly of the canonical PIC. Core promoter elements determine both the position of the TSS and the basal activity of the core promoter [1,3,4]. Core promoter elements have been most extensively characterized in metazoans and include the TATA box, the Initiator (INR), the TFIIB recognition element (BRE) "upstream" and "downstream" (BREu and BREd), the DPE, the MTE, the DCE, the TCT, and others [4–6]. Except for the TATA box and INR-like sequences, similar elements have not been found in yeast promoters [7]. In metazoans, the TATA box is located at −30 bp upstream of the TSS [8] and is bound by the TATA-binding protein (TBP) subunit of the TFIID complex. TFIID is composed of TBP and up to fourteen TBP-associated factors (TAFs) that form stable complexes in metazoans [9–11]; in contrast, TBP in yeast can be recruited to certain promoters independently of TAFs [12,13]. TBP binding to the TATA box is the first step in core promoter recognition and leads to a sharp DNA bend that is important for TFIIB recruitment and, hence, for the assembly and stability of the PIC

[1,3,9,14]. The BRE elements that flank the TATA box in certain promoters are recognized by TFIIB; this further stabilizes the bent TBP-DNA complex via both TFIIB-TBP and TFIIB-DNA interactions [14,15]. The INR is located at the TSS ($+1$) and defines the transcription initiation site independently of, or in synergy with, the TATA box [4]. In contrast, the DCE, MTE and DPE elements are located downstream of the TSS and increase the overall activity of the promoter in synergy with either the TATA box or the INR element [5]. The INR and these downstream core promoter elements are all recognized by different TAFs of the TFIID complex: the INR is bound by TAF1 and TAF2; the DCE is recognized by TAF1; and the MTE and DPE were shown to crosslink to TAF6 and TAF9 [4–6]. Importantly, these elements are not present in all core promoters, and core promoters can be classified according the presence or the absence of these elements (e.g., TATA, TATA/INR, INR, INR/DPE). For instance, the large group of housekeeping gene promoters that are often located within CpG islands do not have recognizable core promoter elements and direct transcription initiation from multiple dispersed TSSs [6]. The TCT element is found at the TSS of a select group of TATA-less genes such as the ribosomal protein genes, and is recognized by the gene-specific TBP-related factor 2 (TRF2) rather than by the canonical TBP/TFIID [16]. TRF2 also binds and stimulates certain DPE-dependent promoters in Drosophila [17]. Often, the members of the metazoan-specific family of TBP-related factors (TRFs) control core promoters of highly-specialized developmental and/or cell type-specific genes that do not depend on the TBP/TFIID basal transcription factor [18–20]. For the vast majority of canonical TBP/TFIID-dependent genes, the core promoter sequence also influences the mechanisms of transcription initiation by RNAPII. This was indicated, for instance, by (i) the differential requirements and functions of TBP and TAFs at TATA-dependent and TATA-less promoters [12,13,21–27]; and (ii) the requirement of accessory core promoter-specific cofactors for the TAF-dependent basal activity of mammalian INR-dependent and DPE-dependent core promoters [28–30] and for the TFIID-dependent synergy of TATA and INR elements [28,29]. Hence, there is no universal mechanism (or general machinery) for specific core promoter recognition and transcription initiation by eukaryotic RNAPII, and this is also true for the canonical TFIID-dependent basal machinery in metazoans, which often relies on auxiliary core promoter-specific cofactors [31].

Besides influencing the basal transcription initiation process and the overall transcription output, increasing evidence also indicates an important regulatory role of the core promoter in the preferential response of specific genes to the regulatory signals of certain activators and enhancers [4,32]. Many Drosophila enhancers were shown recently to preferentially activate a TCT/TRF2-dependent promoter rather than a TFIID-dependent core promoter [33]. However, enhancers and activators can also differentially stimulate canonical TFIID-dependent transcription in a core promoter-dependent manner. Early observations showed that different TATA box sequences have different abilities to convey the activating signals of certain enhancers and activators in mammalian cells [34,35] and in yeast [36–38]. The activation domains of VP16 and SP1 fused to the yeast GAL4 DNA-binding domain were also reported to have different core promoter preferences in mammalian cells, although no strict core promoter selectivity was observed since both VP16 and SP1 could significantly activate all core promoter types tested [39]. In contrast, the upstream D′/Elf-1 sequence element of the mouse *TdT* gene promoter (a TATA-less, INR-containing promoter) was shown to activate selectively INR-containing core promoters (with or without a TATA box) but not a TATA-only core promoter [40,41]. In Drosophila, the promoter preference of certain enhancers has also long been known [42–44]. Promoter competition experiments demonstrated that the Drosophila AE1 and IAB5 enhancers preferentially activated transcription from the TATA-containing *even-skipped* core promoter relative to the TATA-less *white* core promoter. However, there was no strict core promoter selectivity since in the absence of a competing TATA-containing promoter both enhancers could activate the TATA-less *white* promoter [45]. An enhancer-trapping screen in Drosophila

identified three enhancers that selectively activated an INR/DPE core promoter and one enhancer that preferentially activated a TATA/INR core promoter [46]. However, whether the core promoter preference of the TATA/INR-specific enhancer was solely due to the intrinsic function of the TATA box (rather than the INR or the combined/cooperative activities of TATA and INR elements) remained unclear. More recently, investigation of TATA-less DPE-dependent promoters of Drosophila developmental genes identified the enhancer factors *Caudal*, *Relish* and *Dorsal* as activators that preferentially stimulate transcription through the DPE [47–49]. Although each of these activators could also weakly stimulate transcription from a TATA-dependent promoter (TATA/INR), *Caudal* could not activate a TATA promoter containing an upstream BRE (BREu), indicating that core promoter context and cognate basal transcription mechanisms are important for the extent and selectivity of activation by a specific activator [32,47]. Natural TATA-selective activating DNA sequences and cognate TATA-specific activators that do not act through an INR-directed mechanism remained to be identified and characterized.

Here, we found that the natural serum-responsive upstream activating sequence (UAS) of the human β-actin gene (ACTB) selectively activates TATA box-dependent but not INR-dependent transcription via a mechanism that involves the serum response factor (SRF) activator and the TATA-binding/bending activity of TBP in live human cells. In contrast, the TATA-binding/bending activity of TBP is not essential for TFIID-dependent transcription of the endogenous TATA-less *CDKN1A* gene. We further show that activation by the *ACTB* UAS is unidirectional on TATA but bi-directional on TATA/INR core promoters, revealing a stereospecific INR dependency for activation by a natural SRF-responsive UAS. Consistent with these results, computational analyses further indicate that SRF-regulated promoters of the actin cytoskeleton and contractile gene family often contain a TATA box. Our results underscore the regulatory role of the core promoter and support the idea that specific core promoter elements facilitate the coordinated regulation of specific families of genes with related biological functions, and demonstrate the existence of distinct TBP/TFIID-mediated mechanisms of transcription initiation at TATA-dependent and TATA-independent promoters in mammalian cells.

2. Materials and methods

2.1. Plasmid DNAs

The *ACTB* promoter-luciferase construct was previously described [29]. The *ACTB*[+INR] luciferase construct was constructed using the QC-PCR kit (Stratagene) to mutate the transcription start site (TSS) region of *ACTB* construct into a consensus mouse *TdT* initiator using primer: 5′-GGCGGCGCGACGCGCC**ctcattct**CGAGACCGCGTCCGCCCC-3′(mutated sequence in bold) according to manufacturer's instructions. The HSV *TK* luciferase construct was described previously [50]. The HSV *TK*[+INR] was constructed, using the QC-PCR kit (Stratagene) as above, to mutate the TSS region of HSV *TK* into a consensus mouse *TdT* Initiator with the primer: 5′-GTGACGCGTGTGGCCTCC**ctcattct**GCGACCCTGCAGCGACC-3′. The G5-TATA, G5-TATA/INR and G5-INR luciferase constructs were generated by inserting a HindIII DNA fragment containing the five Gal4-binding sites and the core promoter from pG5TdT(-41TATA/+33), pG5TdT(-41TATA/Inr + 33) and pG5TdT(-41Inr/+33) [28] into the HindIII site of pGL3-basic luciferase expression vector (Promega). The mammalian expression vectors for Gal4(1–147), Gal4-SP1(132–243), Gal4-CTF1(399–499) and Gal4-VP16(411–490) were provided by Dr. Nicolas Mermod [51]. The upstream activating sequence of human *ACTB* promoter (-120 to -40 bp), UAS^{ACTB}, was amplified by PCR (primers: 5′-TCTAGT**ggtacc**GCGAAGCCGGTGAGTGAGCG-3′ and 5′-TACATA**ggtacc**GCGGCCGCTCGAGCCATAAAAGGC-3′) and inserted into the KpnI site of TATA, TATA/INR or INR core promoter-luciferase plasmids [29] in the forward or reverse direction to generate UAS^{ACTB}-TATA,

UAS^{ACTB}-TATA/INR, UAS^{ACTB}-INR and their reverse constructs. The mut.CCAAT, mut.CArG, mut.Both, 5xCCAT, 5xCArG and 5xGC constructs were generated by inserting synthetic DNA oligonucleotides into the KpnI site of TATA, TATA/INR and INR core promoter-luciferase plasmids (see all the synthetic oligonucleotide sequences in Supplementary Materials and Methods). The *CDKN1A*-Luciferase reporter was the p21P plasmid described previously [52]. All constructs were confirmed by DNA sequencing.

2.2. Cell culture, transient transfection, luciferase assay, and RNAi

HEK293 cells, HeLa S3 cells, and the derivative HeLa S3 + T210K cell line described previously [22] were cultured in Dulbecco's modified Eagles medium supplied with 10% fetal bovine serum and 5% CO_2 at 37 °C. Transient transfections of HEK293 cells and HeLa S3 cells were performed with Lipofectamine 2000 (Invitrogen), and luciferase assays were performed as described previously [53]. Cells in a 6-well plate were transfected with luciferase reporter plasmids (0.8 μg to 2 μg) and 0.2 μg of pCMV-β-Galactosidase internal control plasmid. Luciferase assays were performed 20 to 48 h after transfection. Luciferase activities were analyzed in triplicates from three or more independent transfections and were normalized to β-Gal activity and to the luciferase activity of cells transfected with the promoter-less pGL3-basic vector (arbitrarily set to 1.0). Control siRNAs were purchased from Ambion (AM4635#1 or AM4637#2). SRF and NFYB siRNAs were previously described: siSRF797 [54] and siNF-YB2 in [55], respectively. The siRNA sequences for siNFYA were 5′-CGUCUAUCAACCAGUUAAUdTdT-3′ and 5′-AUUAACUGGUUGAUAGACGdTdT-3′ for sense and anti-sense strands. The sequences of TBP 3′-UTR siRNA were 5′-GUGACUGUGAGUUGCU CAUdTdT-3′ and 5′-AUGAGCAACUCACAGUCACdTdT-3′ for sense and anti-sense strand. All siRNAs were synthesized by Ambion, except the SP1 siRNA, which was from Cell Signaling (Cat#12104). For TBP knockdown, cells were seeded over night to reach 70% confluency at the time of transfection. 200 nM TBP siRNA was transfected into Hela S3 or Hela S3 + T210K cells at time 0 h. Twelve hours later, 200 nM TBP siRNA was transfected again together with the indicated luciferase constructs. Twenty hours after the second transfection the cells were harvested and half of the cells were used for luciferase assay and the other half for Western blot. Alternatively, to analyze expression of specific endogenous genes, total cellular protein and RNA were isolated from cells transfected with control or TBP siRNAs and analyzed by Western blot and RT-qPCR (see below). For the knockdown of all other factors, 200 nM siRNA was transfected only once into HEK293 cells together with 0.8 μg of *ACTB*-Luc. Forty-eight hours after transfection, half of the cells were used for luciferase assay and the other half for Western blot. Luciferase activities in Fig. 6 were normalized to total cellular protein.

2.3. Antibodies and Western blots

The antibodies SRF(sc-335), NFYB(sc-13045×), NFYA(sc-10779×), SP1(sc-59×), CRSP70(sc-48776) and TAF9/TAF$_{II}$31(sc-1248) were purchased from Santa Cruz Biotechnology. Vinculin (V4505) monoclonal antibody was purchased from Sigma. Rabbit polyclonal antibodies against TBP, TAF4/TAF$_{II}$135 and TAF5/TAF$_{II}$100 were gifts from Robert G. Roeder. Western blots were performed essentially as previously described [53].

2.4. In vitro transcription and primer extension

Transcription reactions with HeLa nuclear extracts and primer extension were described previously [28]. The extension primer for luciferase transcripts was described previously [29]. Total RNA was purified from HEK293 cells using Qiagen RNeasy kit (Cat# 74104). 10 μg of total cellular RNA was used in primer extension analyses as previously described [29]. The X-ray films were scanned with HP

Precisionscan Pro 3.1 scanner and signals were quantitated with NIH ImageJ.

2.5. Reverse transcription and real-time quantitative PCR (RT-qPCR)

Total cell RNA from three biological replicates was purified as above. Reverse transcription and real time PCR were performed with the BioRad iScript™ Reverse Transcription Supermix (#170-8840) and iQ™ SYBR® Green Supermix (#170-8882) per manufacturer's instructions. Real time PCR was performed in triplicates on a BioRad MiniOpticon system. PCR primers (forward and reverse) were *ACTB* mRNA: 5′-TGACGGGGTCACCCACACTGTGCCCA-3′ and 5′-CTAGAAGC ATTTGCGGTGGACGATGGAGGG-3′; *ACTB* pre-mRNA: 5′-GGCACCAC ACCTTCTACAATG-3′ (exon) and 5′-CCACCAGAAGAGGTAGCGGG-3′ (intron); *MYC* mRNA: 5′-CGACGCGGGGAGGCTATTCTGC-3′ and 5′-CCCGCCACCGCCGTCGTTGTCT-3′; and *CDKN1A* mRNA: 5′-GGCAGACC AGCATGACAGATT-3′ and 5′-GCGGATTAGGGCTTCTCTTT-3′. Relative expression of specific mRNAs in siRNA-treated cells was normalized to cells transfected with the control siRNA and was obtained using the Pfaffl method. The normalized means (± S.D.) are presented for three independent experiments.

2.6. Statistics and computational analyses of core promoters

Significant differences ($P < 0.05$) in gene expression analyses were analyzed by two-tailed Student's t-test of at least three independent experiments. Global frequencies of 9010 human core promoters and 7995 conserved orthologous mouse core promoters with experimentally validated transcription start sites were reported previously [56] and used to classify the core promoters of SRF target genes (below) into TATA-containing and TATA-less groups. All SRF target genes were obtained from previous ChIP-seq analyses in mouse fibroblasts [Table 5 in [57]]. The list of SRF target genes of the cytoskeleton-contractile family was obtained from a previous report [Table 1 in [58]]. Significant deviations from the global TATA-containing/TATA-less core promoter frequencies for all SRF target genes and for the specific group of SRF-associated cytoskeleton-contractile genes were analyzed by Chi-square test.

3. Results

3.1. The ACTB gene upstream activating sequence (UAS) stimulates TATA box-dependent but not Initiator-dependent transcription

In an attempt to generate a highly active *ACTB* promoter for reporter gene assays, we introduced by site-directed mutagenesis a perfect Initiator (INR) element at the transcription start site downstream of the TATA box of the *ACTB* promoter in the context of a luciferase reporter plasmid (Fig. 1A). To our surprise the activity of the INR-containing *ACTB* promoter (*ACTB*[+INR]) in transfected HEK293 cells was not stimulated by the INR, whereas the activity of the TATA-containing HSV *TK* promoter was, as expected, increased by the INR (Fig. 1B). Notably, the INR sequence was functional and stimulated transcription in the context of the *ACTB* promoter (*ACTB*[+INR]) in transcription experiments in vitro using HeLa cell nuclear extracts (Fig. 1C). We reasoned that the different behavior of the *ACTB*[+INR] promoter in transfected cells and in vitro could be due to the differential activity of the *ACTB* upstream activating sequence (UAS) and associated activators/coactivators in cells vs. in nuclear extracts (see further below). Hence, these results suggested the possibility that the UAS of *ACTB* and associated sequence-specific activators could be unable to stimulate INR-dependent transcription under more physiological conditions in live cells. This was highly intriguing since, to our knowledge, all activators tested so far in mammalian cells support INR-dependent transcription to some extent. We further investigated this in our cell system by analyzing the ability of distinct activation domains fused to the yeast Gal4 DNA-binding domain to activate three core promoters that differed by only the presence
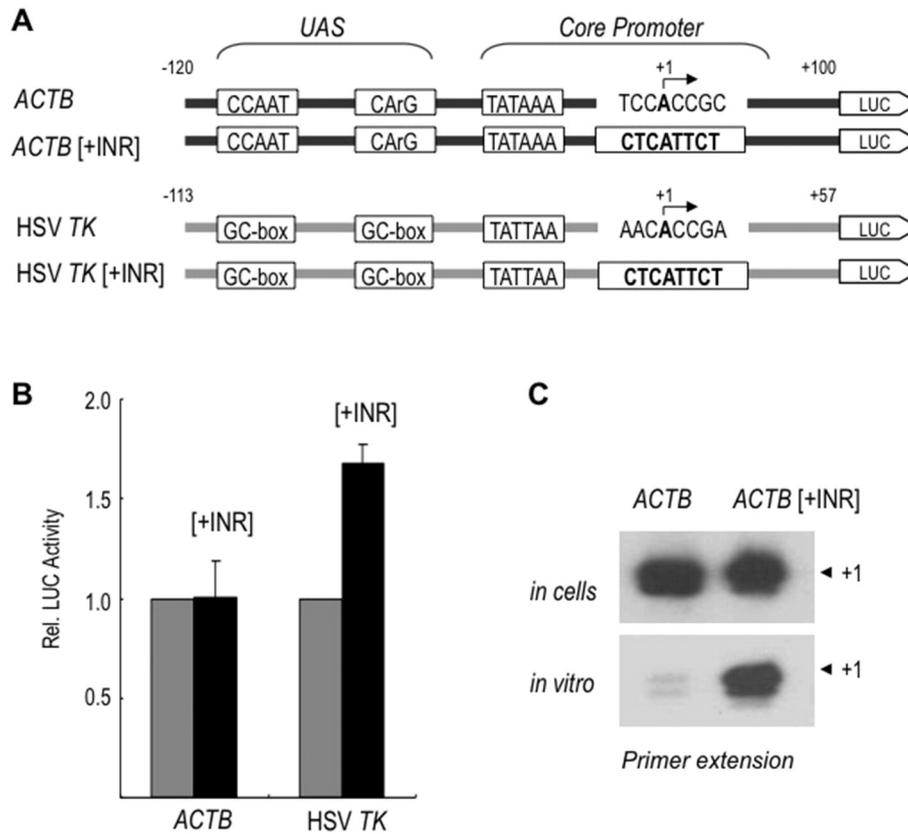
Fig. 1. An INR cannot stimulate transcription from the human *ACTB* promoter in vivo. A. Schematic diagrams of luciferase reporters containing the natural *ACTB* promoter (−120/+100 bp) or the natural HSV *TK* promoter (−113/+57 bp) and derivatives [+INR] containing a consensus INR sequence (CTCATTCT) indicated in bold at the transcription start site (bent arrow; +1). The TATA box sequence of *ACTB* is TATAAA and that of HSV *TK* is TATTAA. The upstream activating sequence (UAS) of *ACTB* contains two essential elements: a CCAAT box and a CArG box. The UAS of HSV *TK* promoter contains two essential GC-box elements. B. Relative luciferase activities of the constructs in panel A transfected into HEK293 cells. Activities of *ACTB*-Luc and HSV-*TK*-Luc lacking an INR (gray bars) were arbitrarily set to 1. Results are the means ± S.D. of three independent experiments. **C.** Primer extension analyses of RNAs transcribed from the *ACTB*-Luc or *ACTB*[+INR]-Luc reporters transfected in HEK293 cells (top panel) or after in vitro transcription in a HeLa nuclear extract (bottom). Correctly initiated transcripts are indicated by "+1".

or absence of a TATA and/or INR element: i.e., TATA-only (T), INR-only (I), and combined TATA/INR (T/I) core promoters (Fig. 2A). Consistent with previous observations, the glutamine-rich activation domain of SP1, the proline-rich activation domain of CTF1, and the acidic activation domain of the viral VP16 activator were all able to activate these three types of core promoters, and the activity of the composite T/I core promoter was always the highest, as expected (Fig. 2B; and Supplemental Fig.S1A). We then substituted the Gal4 UAS sequences with the UAS of *ACTB* (−120/−40), which include a CCAAT box and a CArG box, upstream of the three heterologous core promoters (T, T/I and I; Fig. 3, top) and analyzed their activities similarly. As shown in Fig. 3 (bottom, WT lanes 1–3), the UAS of ACTB did not significantly activate the TATA-less INR-only (I) core promoter and activated TATA-only (T) and composite TATA/INR (T/I) core promoters to a similar extent, as measured by both the relative luciferase activities and the levels of correctly initiated transcripts (+1) by primer extension. Similar results were obtained in HeLa cells indicating that this core promoter selectivity is not cell type-specific (Supplemental Fig. S1B). These results demonstrated the requirement of a TATA box for activation by the UAS of *ACTB* and the inability of these activating sequences and cognate activator/coactivators to stimulate an INR-dependent transcription pathway.

The UAS of *ACTB* contains a CCAAT box and a CArG box, which were shown previously to be both important for the activity of the promoter [59–61]. To determine which of these elements is responsible for the TATA-selectivity and INR-independence of the UAS, we introduced mutations in each of these elements that impair binding of cognate transcription factors NFY and SRF [61] and analyzed the activity and core promoter selectivity of the mutated UAS, as above. As expected,

mutation of either the CCAAT box (mut.CCAAT) or the CArG box (mut.CArG) strongly decreased the activity of the UAS from both TATA-only (T) and composite TATA/INR core promoters (T/I), while the TATA-less INR-only (I) core promoter remained inactive (Fig. 3 bottom). Notably, there was a residual activity of the mut.CCAAT and of the mut.CArG UAS above the basal activity of the promoters containing both elements mutated (mut.Both). Interestingly, although the residual activation by the mut.CCAAT and the mut.CArG UAS was still TATA-dependent (i.e. the TATA-less INR promoter was not stimulated), the INR was now required in addition to the TATA-box for efficient transcription (Fig. 3, bottom panel). To further analyze the intrinsic activity of the CCAAT box and CArG box separately and independently of natural flanking UAS sequences, 5 copies of each of these elements were cloned upstream of the different core promoters (Supplemental Fig.S2). We found that 5 copies of the CArG box by itself were not sufficient for any significant activation, while 5 copies of the CCAAT box activated transcription but were not sufficient for high levels of activation from the TATA-only promoter and required the INR element for increased activation; this was also observed for 5 copies of an SP1/GC-box and for activation by all Gal4 fusion activators, which activated transcription from all promoters as shown above (Fig. 2). This result is also consistent with previous observations that CCAAT boxes bound by NFY can activate both TATA-dependent and TATA-less promoters [62,63]. In vitro transcription experiments with HeLa nuclear extracts and the above mutant UAS constructs confirmed that the CCAAT box (but not the CArG box) is sufficient for a low level of activated transcription and that HeLa nuclear extracts lack a cofactor for TATA-specific activation by the UAS of *ACTB* (Supplemental Fig.S3). Thus, we conclude that it is
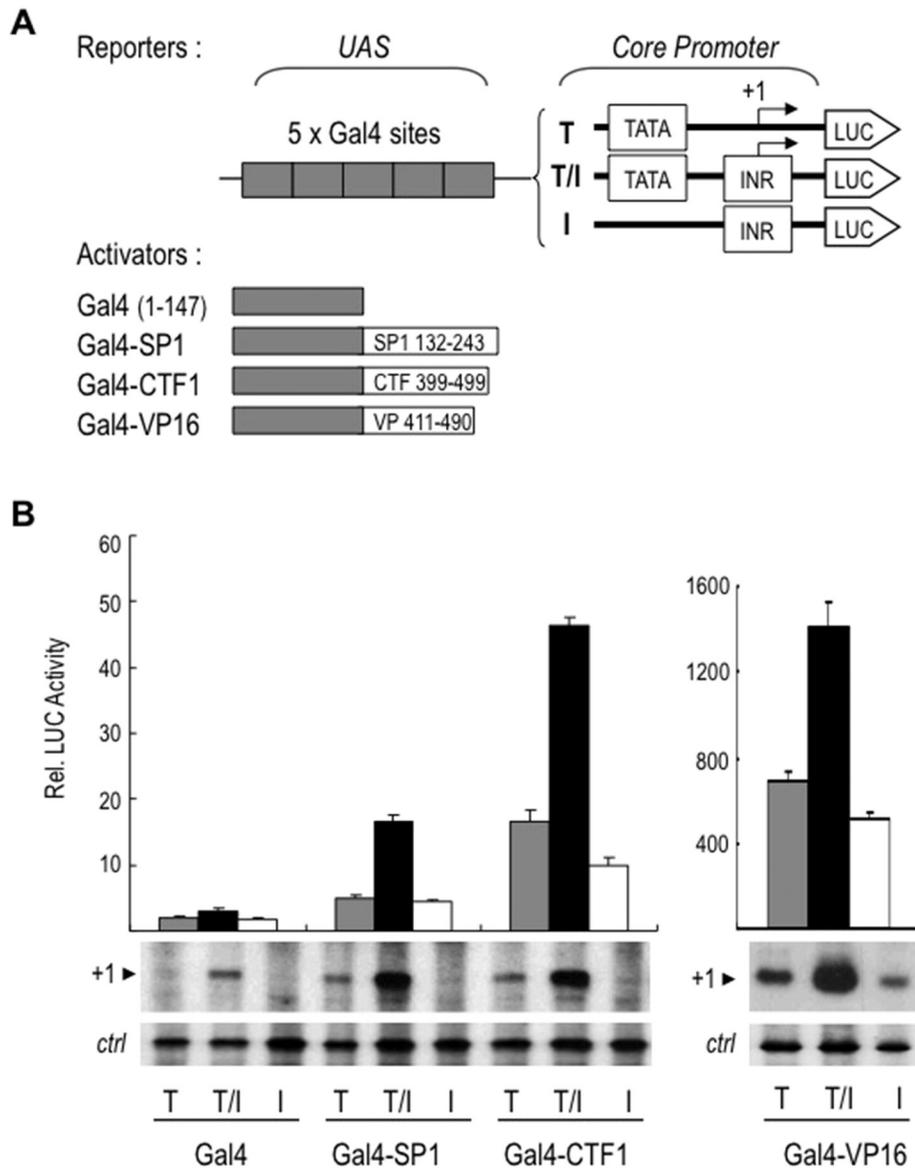
Fig. 2. The INR stimulates transcription activation by different types of activation domains in vivo. A. Schematic diagram of promoter-luciferase reporters and Gal4-fusion activators. Reporters contain five Gal4 binding sites upstream of TATA (T), TATA/INR (T/I) or INR (I) core promoters. Gal4(1–147) is the yeast Gal4 DNA binding domain (amino acids 1–147). The activation domains of SP1 (132–243), CTF1 (399–499) and VP16 (411–490) were fused to Gal4(1–147) to generate Gal4-SP1, Gal4-CTF1 and Gal4-VP16 fusion activators, respectively. B. The reporters above were co-transfected with Gal4(1–147), Gal4-SP1, Gal4-CTF1 or Gal4-VP16 in HEK293 cells. Luciferase activities are relative to the activity of the promoter-less pGL3 luciferase vector, which was arbitrarily set to 1, and are the means ± S.D. of three independent experiments. The basal activities without any activator are shown in Supplemental Fig.S1A. Bottom panels show autoradiograms of representative primer extension analyses indicating correctly initiated transcripts (+1) in transfected cells, including an endogenous cellular transcript used as an internal control (ctrl). The autoradiogram for Gal4-VP16 resulted from a shorter X-ray film exposure.

not the individual activities of each element/box but rather their concerted functions (perhaps also involving additional immediately adjacent sequences) that are responsible for the TATA-specific and INR-independent activity of this UAS.

3.2. Orientation-independent transcription activation by the UAS of ACTB is core promoter-specific and requires an INR element

Our data above indicated that the CCAAT box and CArG box in the UAS of *ACTB* function cooperatively in core promoter selection, possibly as an enhancer-like unit. Indeed it was shown that the CCAAT box and the CArG box need to be on the same face of the DNA helix for UAS function, whereas the activity of the UAS was unaffected by increasing its distance from the TATA box by full or half helical turns [61]. To further address whether the UAS of *ACTB* has enhancer-like properties, we tested its ability to activate transcription when cloned

in both orientations upstream of the different heterologous core promoters. As described above, in its natural "Forward" orientation the UAS activated the TATA and composite TATA/INR core promoters to a similar extent (Fig. 4; Forward, lanes T and T/I respectively) but not the TATA-less INR promoter (Forward, lane I). In contrast, in the "Reverse" orientation, activation of the TATA-only promoter was drastically reduced to ~20%, while efficient activation of the composite TATA/INR core promoter was retained (Fig. 4; Reverse, lanes T and T/I, respectively). We conclude that the UAS of ACTB does not have the orientation-independent properties of enhancers for TATA-only core promoters, but that it does have such bi-directional activity on selected core promoters having both TATA and INR elements. Thus, core promoter-selective activation is not only dictated by the identity and cooperativity of the DNA regulatory elements and cognate sequence-specific activators, but also by their stereo-specific arrangement relative to the core promoter. These observations further suggest the possibility
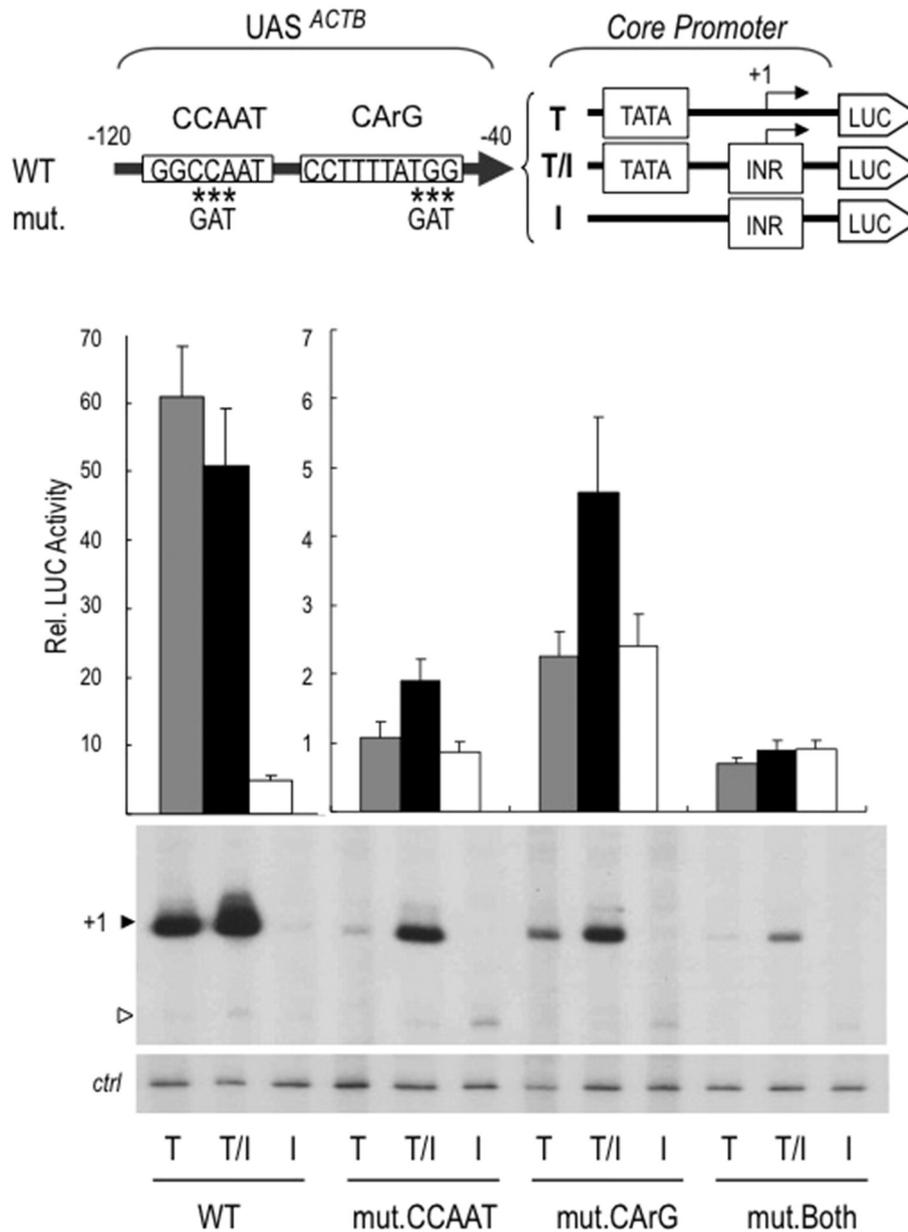
Fig. 3. The UAS of *ACTB* is composed of a CCAAT box and a CArG box that synergistically activate TATA-dependent but not INR-dependent transcription. Top, are diagrams of the reporters containing the wild type (WT) UAS of *ACTB* (− 120 bp to − 40 bp) or mutated (mut.) derivatives upstream of heterologous TATA (T), TATA/INR (T/I) or INR (I) core promoters. The specific nucleotides substituted in the CCAAT-box (mut.CCAAT), CArG-box (mut.CArG) or both boxes (mut.Both) are indicated with asterisks. Bottom, are the relative luciferase activities of wild type and mutant reporters transfected into HEK293 cells (n ≥ 3 independent experiments). The Luc activities are relative to that of the promoter-less pGL3 luciferase vector, which was arbitrarily set to 1. Note the two different scales for the Rel. Luc Activities of WT and mutated *ACTB* UAS reporters. A representative primer extension analysis is also presented showing the correctly initiated transcripts (closed arrowhead; + 1). A low level of incorrectly initiated transcription is seen mostly in the INR (I) lanes (open arrowhead). An endogenous cellular transcript served as internal control (ctrl), as in Fig.3.

that composite core promoters, containing both TATA and INR elements, could be targets for activation by a broader spectrum of regulatory and enhancer sequences.

3.3. The TATA-binding/bending activity of TBP is important in vivo for the TATA-dependent promoters of ACTB and MYC but not for the TATA-independent CDKN1A promoter

Having established the requirement of the TATA box for the activity of the UAS of *ACTB* in transfection/reporter gene assays, we then investigated the importance of such TATA box-dependent pathway for transcription of the endogenous *ACTB* gene in human cells. We used the HeLa S3 cell line and a derivative (HeLa S3 + T210K) that expresses

similar physiologic levels of both the endogenous wild type TBP and an ectopic mutant TBP (TBP-T210K) carrying the T210K amino acid substitution within its TATA-binding domain [22]. The T210K substitution (equivalent to the T112K mutation in yeast TBP) is the most radical TBP mutation reported in terms of abolishing the TATA-specific DNA-binding activity of TBP [27,64–66], and also prevents TBP-mediated DNA bending [67], but does not affect TBP interaction with TAFs and with other components of the basal transcription machinery [22,66]. To test the requirement of TBP and of the TBP-TATA interaction for endogenous *ACTB* gene transcription in HeLa cells, TBP expression was knocked down by transient transfection with a specific siRNA that targets the 3′-UTR of human *TBP* mRNA. This siRNA efficiently and selectively reduced endogenous TBP protein levels in wild type HeLa S3
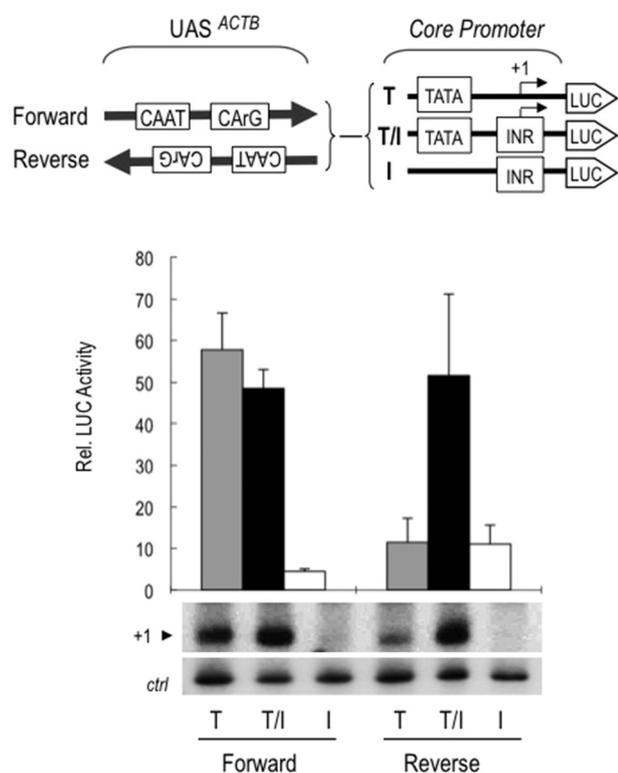
Fig. 4. Transcription activation by the *ACTB* UAS is unidirectional on TATA but bidirectional on a composite TATA/INR core promoter. The top part shows the luciferase reporters containing the UAS of *ACTB* (−120 to −40 bp) inserted upstream of heterologous TATA, TATA/INR or INR core promoters in forward or reverse orientation. The lower part shows the luciferase activities of the above reporters in transfected HEK293 cells. Luciferase activities are the means of three independent experiments. Primer extension analyses of correctly initiated Luc transcripts (+1), including an internal control (ctrl) transcript (as in Figs. 2 and 3), are shown at the bottom.

cells and in HeLa S3 + T210K cells (Fig. 5A, lanes 3 and 6, endo.WT; see also Supplemental Fig.S4) but did not affect the expression of the ectopic mutant TBP-T210K (Fig. 5A, lane 6, T210K). This provided us with an assay to determine whether replacement of endogenous TBP with TBP-T210K affects transcription of *ACTB* and other genes. After knocking down TBP in wild type cells we analyzed both steady-state *ACTB* mRNA and pre-mRNA levels by reverse transcription and quantitative real-time PCR (RT-qPCR). *ACTB* steady state mRNA levels were not affected by the transient knockdown of TBP, probably due to the high abundance and/or stability of *ACTB* mRNA; in contrast, *ACTB* pre-mRNA levels, which better reflect de novo transcription, were reduced significantly and to similar extents in both wild type HeLa S3 cells and in HeLa S3 + T210K cells (Fig. 5B, *ACTB* black and gray bars, respectively). Thus, the TBP-T210K mutant protein could not rescue transcription of *ACTB* in cells depleted of endogenous TBP, confirming the requirement of the TBP-TATA interaction in vivo for endogenous *ACTB* gene transcription. Similar results were also obtained for the TATA-containing *MYC* gene. In contrast, TBP-dependent transcription of the TATA-less *CDKN1A* gene was rescued by the mutant TBP-T210K (Fig. 5B, *MYC* vs. *CDKN1A*). Consistent with these results, promoter-luciferase assays in TBP knockdown cells indicated that the activities of the wild type *ACTB* promoter and of its UAS cloned upstream of a heterologous TATA-containing core promoter could not be rescued by the TBP-T210K mutant, while the activity of the *CDKN1A* promoter was rescued by TBP-T210K (Fig. 5C). These results demonstrate that the in vivo TBP-TATA interaction is important for activation of the *ACTB* and other TATA-dependent promoters but is dispensable for TBP-dependent transcription from at least some TATA-less (or TATA-independent) promoters in mammalian cells.

3.4. Serum response factor (SRF)-dependent actin cytoskeleton and contractile genes tend to have TATA-containing core promoters

Transcription activators that stimulate selectively TATA-dependent but not INR-directed transcription have not yet been described. The identification of the UAS of *ACTB* as a TATA-specific activating unit suggested a role for the DNA-binding factors SRF and NFY, which were shown originally to bind, respectively, the CArG box and CCAAT box sequences of this UAS in vitro [61]. These sequences are also bound in vivo by SRF and NFY in a variety of human cell lines, including in HeLa S3 cells, as indicated by the ENCODE transcription factor ChIP-seq datasets at UCSC Genome Browser. In addition, a GC box partially overlapping with the CCAAT box is also bound by SP1 in a subset of cell lines (ENCODE data not shown). While the binding of SRF to the CArG box and SRF-mediated activation of mammalian *ACTB* genes in response to serum-induced signaling is now well established [57,58], the role of NFY or SP1 is still unclear. As expected, siRNA-mediated knockdown of SRF reduced the activity of the *ACTB* promoter (Fig .6A and B). However, knockdown of two NFY subunits (NFYA and NFYB) or SP1 did not inhibit the *ACTB* promoter (Fig. 6A and B), while a control 5xGC-box reporter was inhibited (data not shown). Surprisingly, knockdown of NFYA but not NFYB increased *ACTB* promoter activity, suggesting a possible repressive role of NFYA. Thus, SRF and a still elusive cooperating factor(s) that binds this GC/CAAT box (Fig. 3) and appears to be distinct from conventional NFY or SP1 is required for the TATA-selective activity of the UAS of *ACTB*.

To determine whether SRF is more generally involved in regulation of TATA box-dependent transcription, we analyzed the promoters of all SRF target genes identified by ChIP-seq in mouse fibroblasts [57]. We only considered genes with experimentally validated transcription start sites and well characterized core promoters classified previously as either TATA-containing or TATA-less [56]. As shown in Table 1, the global frequency of TATA-containing core promoters in mouse and human genes was found to be only ~16% [56]. In contrast, by the same criteria about 22% of all SRF target gene promoters in mouse fibroblasts contain a TATA box. Although this moderate enrichment was statistically significant ($P = 0.0024$), it also suggested that only a fraction of SRF-regulated genes might be TATA-dependent, as expected from the fact that SRF cooperates with other factors to elicit the TATA selectivity at the UAS of *ACTB*. Hence, we then focused our analysis more specifically on the promoters of the actin-related group of SRF-regulated genes having cytoskeleton or contractile functions since SRF is a well established master regulator of the actin cytoskeleton [58]. Interestingly, a larger fraction, i.e., about 40% (38% in human and 43% in mouse) of the core promoters for these cytoskeleton-contractile genes have been classified as TATA-containing (Table 1). This is a substantial and statistically very significant enrichment compared to the low global frequency of TATA-containing promoters in the genome ($P < 10^{-7}$). Altogether these results show that SRF has a regulatory bias for TATA-containing promoters genome-wide and, more specifically, for its targeted group of genes encoding the actin cytoskeleton and contractile apparatus.

4. Discussion

The importance of core promoter DNA elements in regulation of transcription by upstream promoter-bound activators and distal enhancers has long been proposed. However, only few natural cis-acting DNA sequences and trans-acting factors with core promoter selectivity have been reported and core promoter-specific mechanisms for transcription initiation by the canonical TBP/TFIID-dependent basal transcription machinery have remained poorly documented in vivo (see Introduction). Here, we have demonstrated the importance of a TATA box, and of the TATA-binding/bending activity of TBP, for in vivo activation by the SRF-dependent UAS of the human *ACTB* gene and the inability of these activating sequences to stimulate transcription through an INR-dependent mechanism. To our knowledge, this is the first report
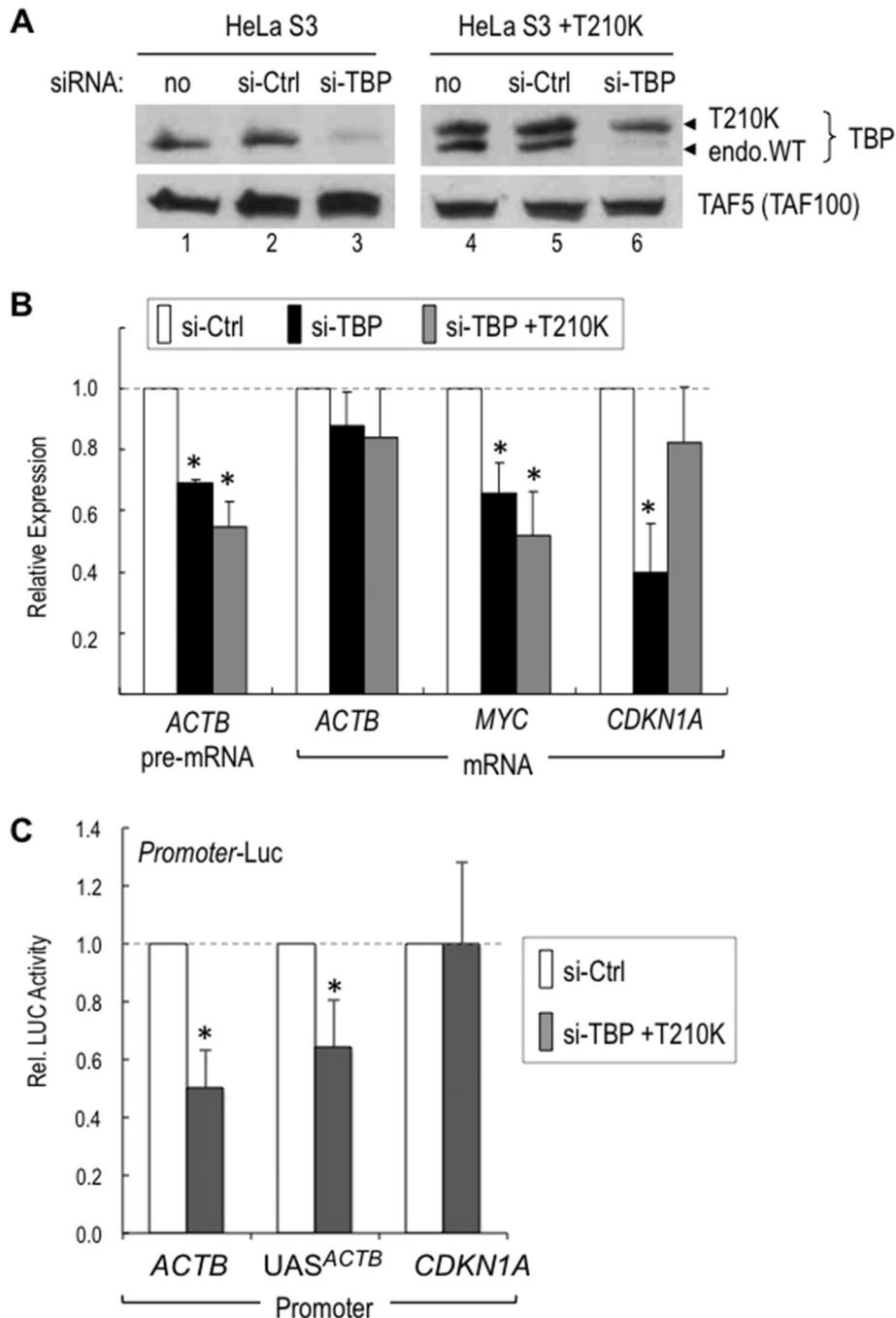
Fig. 5. The TATA-binding activity of TBP is required in vivo for the TATA-containing *ACTB* and *MYC* promoters but not for the TATA-less *CDKN1A* promoter. A. Western blot analyses of endogenous TBP knockdown in Hela S3 and HeLa S3 + T210K cell lines. The top stripe was probed with a TBP antibody. Endogenous TBP wild type (endo.WT) and HA-tagged TBP-T210K mutant (T210K) proteins are indicated with arrowheads. The bottom stripe was probed with a TAF5/TAF100 antibody. No siRNA (lanes 1 and 4), a control siRNA (lanes 2 and 5) or the TBP 3′-UTR siRNA were transfected into both cell lines, as indicated. TBP knockdown under the conditions used (see Materials and Methods) did not affect expression of TAFs or other proteins (see also Supplemental Fig. S4). B. Expression levels of *ACTB* pre-mRNA, and *ACTB*, *MYC* and *CDKN1A* mRNAs were quantitated by RT-qPCR in HeLa S3 (black bars) and HeLa S3 + T210K (gray bars) cells after transfection with the TBP siRNA (si-TBP) and are relative to the levels of expression of each gene transcript in cells transfected with the control siRNA (si-Ctrl; white bars), arbitrarily set to 1. The asterisks indicate significant differences relative to si-Ctrl transfected cells ($P < 0.05$). **C.** Relative activities of promoter-luciferase constructs containing the natural *ACTB* or CDKN1A promoters or UAS-ACTB-TATA. Luciferase reporters were transfected into Hela S3 + T210K cells in the conditions of mock or knockdown of endogenous TBP. Luciferase activity was averaged from three independent experiments. Luciferase activity from si-Ctrl samples was arbitrarily set to 1.

of a TATA box-selective activating sequence and cognate activator. Our results further indicate that it is not simply the individual activities of the SRF-bound CArG box or the adjacent essential GC/CAAT box but rather their concerted functions, as a unit, that are responsible for the TATA-specific and INR-independent activity of this UAS. However, the GC/CAAT box-associated factors that cooperate with SRF remain to be identified. Indeed, although SP1 and NFY factors bind this GC/CAAT box in various cell types (from data of the ENCODE project), our RNAi

analyses suggest that neither SP1 nor the conventional NFY trimeric complex (composed of the NFYA, B and C subunits) is required for activation by this UAS. Instead, knockdown of NFYA stimulated the activity of the UAS, while knockdown of NFYB had no apparent effect, suggesting a repressive function of NFYA and dispensability of NFYB at the *ACTB* UAS. This would be consistent with the reported dual positive and negative transcription activities of NFY [68], and with previous observations indicating that the activities of NFYA and NFYB do not always
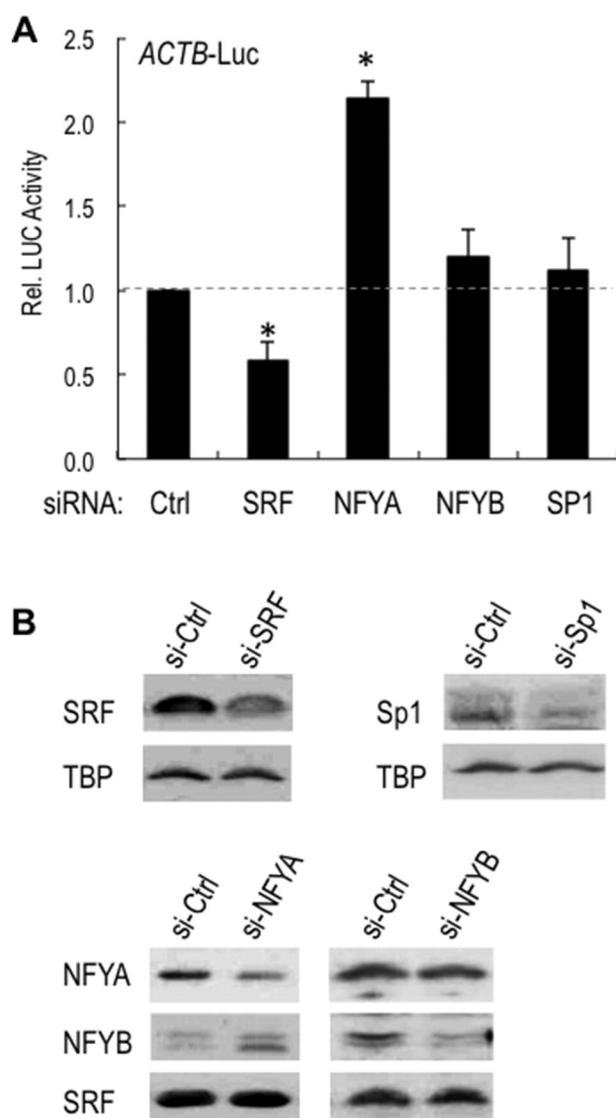
Fig. 6. *ACTB* promoter activity requires SRF but not NFY or SP1. A. Relative luciferase activity from the natural *ACTB* promoter-luciferase construct (see Fig.1A) in cells transfected with control (Ctrl), SRF, NFYA, NFYB or SP1 siRNAs. Luciferase activity was normalized to total protein. Luciferase activity in cells transfected with the siRNA Ctrl was arbitrarily set to 1. Results are the means from three independent experiments. Asterisks indicate significant differences (*P* < 0.05) relative to Ctrl. B. Western blot analysis of SRF, NFYA, NFYB and SP1 proteins in extracts of cells transfected with the control siRNA (si-Ctrl) and specific siRNAs as indicated.

Table 1

Frequencies of core promoter types for mammalian SRF target genes.

Core promoter type	All	TATA-containing	TATA-less
Global frequencies:			
All Characterized[a]			
Human	9010 (100%)	1483 (16.5%)	7527 (83.5%)
Mouse	7995 (100%)	1304 (16.3%)	6691 (83.7%)
SRF target genes (Mouse): All[b]	476 (100%)	103 (21.7%)*	373 (78.3%)*
SRF target genes:			
Cytoskeleton-contractile[c]			
Human	97 (100%)	37 (38.1%)**	60 (61.9%)**
Mouse	82 (100%)	35 (42.7%)***	47 (57.3%)***

Asterisks indicate significant deviation from global genomic frequencies analyzed by Chi-square test: *P = 2.4 E-3; **P = 2.7 E-8; ***P = 4.4 E-10.

[a] Global frequencies of core promoter types for human and mouse orthologous genes with experimentally validated transcription start sites [56].

[b] All SRF target genes identified by ChIP-seq in mouse fibroblasts [57] with characterized core promoters in Jin et al. [56].

[c] Cytoskeleton-contractile SRF target genes (from [58]) with characterized core promoters in Jin et al. [56].

overlap since these two NFY subunits are required for regulation of not only common but also distinct target genes, and have different cellular functions [55,69].

Although activation by the *ACTB* UAS is not influenced by the presence or the absence of an INR element in the core promoter, we found that an INR is required for efficient activation when the UAS is in the non-native reverse orientation relative to the core promoter. Hence, it is not only the nature of the activating sequences and cognate DNA-binding factors but also their stereospecific alignment relative to the core promoter that influences the requirement of specific core promoter elements for efficient activation. These observations have implications for the functional definitions of enhancer sequences, which are characterized by their ability to activate transcription of heterologous promoters in a position- and orientation-independent manner. The *ACTB* UAS was shown previously to activate transcription of heterologous TATA-containing promoters at varying and long distances from the promoter [61]. We now show that if the core promoter also contains

an INR, the UAS of *ACTB* can function in an orientation-independent manner as well. Thus, UAS sequences such as the one in *ACTB* may function as enhancers for only certain core promoters, and composite core promoters that contain both TATA and INR elements could be targets for activation by a broader spectrum of regulatory and enhancer sequences. It is also interesting to note that more than half of mammalian promoters are bi-directional and drive divergent transcription from sense and upstream anti-sense core promoters that are separated and controlled by a common UAS sequence [70,71]. We propose that bidirectional transcription at those promoters is dependent on the compatibility of the shared UAS and cognate activators/coactivators with both the sense and upstream antisense core promoter elements.

SRF binds thousands of sites in the mammalian genome and activates a variety of genes in response to serum-induced intracellular Rho-actin or Ras signaling pathways, and functions via concerted interactions with signal-dependent cofactors (MRTFs or TCFs) and a variety of other cooperating factors that interact with adjacent DNA motifs. Hence, the mechanisms of SRF-dependent activation are gene-/context-specific [57,58]. Accordingly, we found that the core promoter structure of mammalian SRF target genes is also diverse, albeit with a moderate bias for TATA-containing promoters when SRF target genes are analyzed globally. However, SRF is a well established master regulator of the specific family of genes encoding the actin cytoskeleton and contractile apparatus [58], and we found that ~40% of the core promoters for these genes contain a TATA box, which is a significant enrichment compared to the low overall frequency of TATA-containing promoters in human and mouse genomes (Table 1). These results suggest that a TATA box-dependent transcription mechanism is essential for SRF regulation of the actin cytoskeleton/contractile-related gene family and support the idea that core promoter elements play a role in the coordinated regulation of specific families of functionally related genes or genes networks [6,32].

We demonstrated that the native *ACTB* gene and a heterologous TATA-containing promoter controlled by the SRF-dependent UAS of *ACTB* require TBP and its TATA-binding/bending activity in vivo. During these analyses we found that the TATA-containing *MYC* gene behaved similarly but that the TATA-less *CDKN1A* gene functioned differently. Native *CDKN1A* and an artificial *CDKN1A* promoter-luciferase construct were transcribed in a TBP-dependent manner in human cells but did not rely on the TATA-binding/bending activity of TBP (which was inactivated by the T210K substitution in TBP). Similarly, it was shown recently in yeast that several TBP mutants defective in TATA binding were unable to sustain expression of the TATA-containing *HIS4* and *SNZ1* genes but could direct transcription of the ribosomal protein genes *RPS5* and *RPL5* that do not have a consensus TATA element [27]. These findings together with previous evidence indicating that TBP can be recruited to promoters via alternative TAF-dependent mechanisms

[21–23,25,26,37,38] suggest that the in vivo TBP-TATA interaction (including TBP-induced DNA bending) is critical for activation of TATA-dependent promoters but is dispensable for TBP-dependent transcription from at least some TATA-less (or TATA-independent) promoters in all eukaryotes. Hence, the canonical TBP-TATA binding/bending mechanism for PIC assembly characterized on selected TATA-containing promoters in vitro is not universal in vivo. It will be important to determine what fraction of yeast and metazoan genomes do in fact depend on the TATA-binding/bending activity of TBP, especially since most metazoan promoters lack a TATA-like element.

Transparency document

The Transparency document associated with this article can be found, in the online version.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbagrm.2016.01.005.

References

[1] R.G. Roeder, The role of general initiation factors in transcription by RNA polymerase II, Trends Biochem. Sci. 21 (1996) 327–335.
[2] R.G. Roeder, Transcriptional regulation and the role of diverse coactivators in animal cells, FEBS Lett. 579 (2005) 909–915.
[3] S. Hahn, Structure and mechanism of the RNA polymerase II transcription machinery, Nat. Struct. Mol. Biol. 11 (2004) 394–403.
[4] S.T. Smale, J.T. Kadonaga, The RNA polymerase II core promoter, Annu. Rev. Biochem. 72 (2003) 449–479.
[5] T. Juven-Gershon, J.Y. Hsu, J.W. Theisen, J.T. Kadonaga, The RNA polymerase II core promoter — the gateway to transcription, Curr. Opin. Cell Biol. 20 (2008) 253–259.
[6] J.T. Kadonaga, Perspectives on the RNA polymerase II core promoter, Wiley Interdiscip. Rev. Dev. Biol. 1 (2012) 40–51.
[7] C. Yang, E. Bolotin, T. Jiang, F.M. Sladek, E. Martinez, Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters, Gene 389 (2007) 52–65.
[8] M.L. Goldberg, Sequence Analysis of Drosophila Histone Genes, Stanford University, California, 1979.
[9] E. Martinez, Multi-protein complexes in eukaryotic gene transcription, Plant Mol. Biol. 50 (2002) 925–947.
[10] G. Papai, P.A. Weil, P. Schultz, New insights into the function of transcription factor TFIID from recent structural studies, Curr. Opin. Genet. Dev. 21 (2011) 219–224.
[11] L. Tora, A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription, Genes Dev. 16 (2002) 673–675.
[12] X.Y. Li, S.R. Bhaumik, M.R. Green, Distinct classes of yeast promoters revealed by differential TAF recruitment, Science 288 (2000) 1242–1244.
[13] L. Kuras, P. Kosa, M. Mencia, K. Struhl, TAF-containing and TAF-independent forms of transcriptionally active TBP in vivo, Science 288 (2000) 1244–1248.
[14] T. Lagrange, A.N. Kapanidis, H. Tang, D. Reinberg, R.H. Ebright, New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB, Genes Dev. 12 (1998) 34–44.
[15] W. Deng, S.G. Roberts, A core promoter element downstream of the TATA box that is recognized by TFIIB, Genes Dev. 19 (2005) 2418–2423.
[16] Y.L. Wang, S.H. Duttke, K. Chen, J. Johnston, G.A. Kassavetis, J. Zeitlinger, J.T. Kadonaga, TRF2, but not TBP, mediates the transcription of ribosomal protein genes, Genes Dev. 28 (2014) 1550–1555.
[17] A. Kedmi, Y. Zehavi, Y. Glick, Y. Orenstein, D. Ideses, C. Wachtel, T. Doniger, H. Waldman Ben-Asher, N. Muster, J. Thompson, S. Anderson, D. Avrahami, J.R. Yates 3rd, R. Shamir, D. Gerber, T. Juven-Gershon, Drosophila TRF2 is a preferential core promoter regulator, Genes Dev. 28 (2014) 2163–2174.
[18] L. Xiao, M. Kim, J. DeJong, Developmental and cell type-specific regulation of core promoter transcription factors in germ cells of frogs and mice, Gene Expr. Patterns 6 (2006) 409–419.
[19] J. DeJong, Basic mechanisms for the control of germ cell gene expression, Gene 366 (2006) 39–50.
[20] J.A. Goodrich, R. Tjian, Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation, Nat. Rev. Genet. 11 (2010) 549–558.
[21] E. Martinez, C.M. Chiang, H. Ge, R.G. Roeder, TATA-binding protein-associated factor(s) in TFIID function through the initiator to direct basal transcription from a TATA-less class II promoter, EMBO J. 13 (1994) 3115–3126.
[22] E. Martinez, Q. Zhou, N.D. L'Etoile, T. Oelgeschlager, A.J. Berk, R.G. Roeder, Core promoter-specific function of a mutant transcription factor TFIID defective in TATA-box binding, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 11864–11868.
[23] C.P. Verrijzer, J.L. Chen, K. Yokomori, R. Tjian, Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II, Cell 81 (1995) 1115–1125.
[24] S.K. Hansen, R. Tjian, TAFs and TFIIA mediate differential utilization of the tandem Adh promoters, Cell 82 (1995) 565–575.
[25] J. Kaufmann, S.T. Smale, Direct recognition of initiator elements by a component of the transcription factor IID complex, Genes Dev. 8 (1994) 821–829.
[26] S.M. Lauberth, T. Nakayama, X. Wu, A.L. Ferris, Z. Tang, S.H. Hughes, R.G. Roeder, H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation, Cell 152 (2013) 1021–1036.
[27] I. Kamenova, L. Warfield, S. Hahn, Mutations on the DNA binding surface of TBP discriminate between yeast TATA and TATA-less gene transcription, Mol. Cell. Biol. 34 (2014) 2929–2943.
[28] E. Martinez, H. Ge, Y. Tao, C.X. Yuan, V. Palhan, R.G. Roeder, Novel cofactors and TFIIA mediate functional core promoter selectivity by the human TAFII150-containing TFIID complex, Mol. Cell. Biol. 18 (1998) 6571–6583.
[29] M. Xu, P. Sharma, S. Pan, S. Malik, R.G. Roeder, E. Martinez, Core promoter-selective function of HMGA1 and Mediator in Initiator-dependent transcription, Genes Dev. 25 (2011) 2513–2524.
[30] B.A. Lewis, R.J. Sims 3rd, W.S. Lane, D. Reinberg, Functional characterization of core promoter elements: DPE-specific transcription requires the protein kinase CK2 and the PC4 coactivator, Mol. Cell 18 (2005) 471–481.
[31] E. Martinez, Core promoter-selective coregulators of transcription by RNA polymerase II, Transcription 3 (2012) 295–299.
[32] T. Juven-Gershon, J.T. Kadonaga, Regulation of gene expression via the core promoter and the basal transcriptional machinery, Dev. Biol. 339 (2010) 225–229.
[33] M.A. Zabidi, C.D. Arnold, K. Schernhuber, M. Pagani, M. Rath, O. Frank, A. Stark, Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation, Nature 518 (2015) 556–559.
[34] M.C. Simon, T.M. Fisch, B.J. Benecke, J.R. Nevins, N. Heintz, Definition of multiple, functionally distinct TATA elements, one of which is a target in the hsp70 promoter for E1A regulation, Cell 52 (1988) 723–729.
[35] F.C. Wefald, B.H. Devlin, R.S. Williams, Functional heterogeneity of mammalian TATA-box sequences revealed by interaction with a cell-specific enhancer, Nature 344 (1990) 260–262.
[36] V. Iyer, K. Struhl, Mechanism of differential utilization of the his3 TR and TC TATA elements, Mol. Cell. Biol. 15 (1995) 7059–7066.
[37] X.Y. Li, S.R. Bhaumik, X. Zhu, L. Li, W.C. Shen, B.L. Dixit, M.R. Green, Selective recruitment of TAFs by yeast upstream activating sequences. Implications for eukaryotic promoter structure, Curr. Biol. 12 (2002) 1240–1244.
[38] M. Mencia, Z. Moqtaderi, J.V. Geisberg, L. Kuras, K. Struhl, Activator-specific recruitment of TFIID and regulation of ribosomal protein genes in yeast, Mol. Cell 9 (2002) 823–833.
[39] K.H. Emami, W.W. Navarre, S.T. Smale, Core promoter specificities of the Sp1 and VP16 transcriptional activation domains, Mol. Cell. Biol. 15 (1995) 5906–5916.
[40] I.P. Garraway, K. Semple, S.T. Smale, Transcription of the lymphocyte-specific terminal deoxynucleotidyltransferase gene requires a specific core promoter structure, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 4336–4341.
[41] P. Ernst, K. Hahm, L. Trinh, J.N. Davis, M.F. Roussel, C.W. Turck, S.T. Smale, A potential role for Elf-1 in terminal transferase gene regulation, Mol. Cell. Biol. 16 (1996) 6121–6131.
[42] X. Li, M. Noll, Compatibility between enhancers and promoters determines the transcriptional specificity of gooseberry and gooseberry neuro in the Drosophila embryo, EMBO J. 13 (1994) 400–406.
[43] C. Merli, D.E. Bergstrom, J.A. Cygan, R.K. Blackman, Promoter specificity mediates the independent regulation of neighboring genes, Genes Dev. 10 (1996) 1260–1270.
[44] V. Corbin, T. Maniatis, The role of specific enhancer-promoter interactions in the Drosophila Adh promoter switch, Genes Dev. 3 (1989) (2191–-2120).
[45] S. Ohtsuki, M. Levine, H.N. Cai, Different core promoters possess distinct regulatory activities in the Drosophila embryo, Genes Dev. 12 (1998) 547–556.
[46] J.E. Butler, J.T. Kadonaga, Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs, Genes Dev. 15 (2001) 2515–2519.
[47] T. Juven-Gershon, J.Y. Hsu, J.T. Kadonaga, Caudal, a key developmental regulator, is a DPE-specific transcriptional factor, Genes Dev. 22 (2008) 2823–2830.
[48] Y. Zehavi, O. Kuznetsov, A. Ovadia-Shochat, T. Juven-Gershon, Core promoter functions in the regulation of gene expression of Drosophila dorsal target genes, J. Biol. Chem. 289 (2014) 11993–12004.
[49] H. Shir-Shapira, J. Sharabany, M. Filderman, D. Ideses, A. Ovadia-Shochat, M. Mannervik, T. Juven-Gershon, Structure–function analysis of the Drosophila melanogaster caudal transcription factor provides insights into core promoter-preferential activation, J. Biol. Chem. 290 (2015) 17293–17305.
[50] Y.L. Wang, F. Faiola, M. Xu, S. Pan, E. Martinez, Human ATAC is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein, J. Biol. Chem. 283 (2008) 33808–33815.

[51] R. Pankiewicz, Y. Karlen, M.O. Imhof, N. Mermod, Reversal of the silencing of tetracycline-controlled genes requires the coordinate action of distinctly acting transcription factors, J. Gene Med. 7 (2005) 117–132.

[52] M.B. Datto, Y. Yu, X.F. Wang, Functional analysis of the transforming growth factor beta responsive elements in the WAF1/Cip1/p21 promoter, J. Biol. Chem. 270 (1995) 28623–28628.

[53] F. Faiola, X. Liu, S. Lo, S. Pan, K. Zhang, E. Lymar, A. Farina, E. Martinez, Dual regulation of c-Myc by p300 via acetylation-dependent control of Myc protein turnover and coactivation of Myc-induced transcription, Mol. Cell. Biol. 25 (2005) 10220–10234.

[54] D. Werth, G. Grassi, N. Konjer, B. Dapas, R. Farra, C. Giansante, R. Kandolf, G. Guarnieri, A. Nordheim, O. Heidenreich, Proliferation of human primary vascular smooth muscle cells depends on serum response factor, Eur. J. Cell Biol. 89 (2010) 216–224.

[55] P. Benatti, V. Basile, D. Merico, L.I. Fantoni, E. Tagliafico, C. Imbriano, A balance between NF-Y and p53 governs the pro- and anti-apoptotic transcriptional response, Nucleic Acids Res. 36 (2008) 1415–1428.

[56] V.X. Jin, G.A. Singer, F.J. Agosto-Perez, S. Liyanarachchi, R.V. Davuluri, Genome-wide analysis of core promoter elements from conserved human and mouse orthologous pairs, BMC Bioinformatics 7 (2006) 114.

[57] C. Esnault, A. Stewart, F. Gualdrini, P. East, S. Horswell, N. Matthews, R. Treisman, Rho-actin signaling to the MRTF coactivators dominates the immediate transcriptional response to serum in fibroblasts, Genes Dev. 28 (2014) 943–958.

[58] J.M. Miano, X. Long, K. Fujiwara, Serum response factor: master regulator of the actin cytoskeleton and contractile apparatus, Am. J. Physiol. Cell Physiol. 292 (2007) C70–C81.

[59] S.Y. Ng, P. Gunning, S.H. Liu, J. Leavitt, L. Kedes, Regulation of the human beta-actin promoter by upstream and intron domains, Nucleic Acids Res. 17 (1989) 601–615.

[60] R.M. Frederickson, M.R. Micheau, A. Iwamoto, N.G. Miyamoto, 5′ flanking and first intron sequences of the human beta-actin gene required for efficient promoter activity, Nucleic Acids Res. 17 (1989) 253–270.

[61] S.L. Danilition, R.M. Frederickson, C.Y. Taylor, N.G. Miyamoto, Transcription factor binding and spacing constraints in the human beta-actin proximal promoter, Nucleic Acids Res. 19 (1991) 6913–6922.

[62] D. Dolfini, F. Zambelli, G. Pavesi, R. Mantovani, A perspective of promoter architecture from the CCAAT box, Cell Cycle 8 (2009) 4127–4137.

[63] E. Martinez, The NF-Y angle to the CCAAT's tale, Cell Cycle 9 (2010) 642–643.

[64] M.A. Choukrallah, D. Kobi, I. Martianov, W.W. Pijnappel, N. Mischerikow, T. Ye, A.J. Heck, H.T. Timmers, I. Davidson, Interconversion between active and inactive TATA-binding protein transcription complexes in the mouse genome, Nucleic Acids Res. 40 (2012) 1446–1459.

[65] P. Reddy, S. Hahn, Dominant negative mutations in yeast TFIID define a bipartite DNA-binding region, Cell 65 (1991) 349–357.

[66] M.C. Schultz, R.H. Reeder, S. Hahn, Variants of the TATA-binding protein can distinguish subsets of RNA polymerase I, II, and III promoters, Cell 69 (1992) 697–702.

[67] X. Zhao, W. Herr, A regulated two-step mechanism of TBP binding to DNA: a solvent-exposed surface of TBP inhibits TATA box recognition, Cell 108 (2002) 615–627.

[68] M. Ceribelli, D. Dolfini, D. Merico, R. Gatta, A.M. Vigano, G. Pavesi, R. Mantovani, The histone-like NF-Y is a bifunctional transcription factor, Mol. Cell. Biol. 28 (2008) 2047–2058.

[69] P. Benatti, D. Dolfini, A. Vigano, M. Ravo, A. Weisz, C. Imbriano, Specific inhibition of NF-Y subunits triggers different cell proliferation defects, Nucleic Acids Res. 39 (2011) 5356–5368.

[70] B.S. Scruggs, D.A. Gilchrist, S. Nechaev, G.W. Muse, A. Burkholder, D.C. Fargo, K. Adelman, Bidirectional transcription arises from two distinct hubs of transcription factor binding and active chromatin, Mol. Cell 58 (2015) 1101–1112.

[71] S.H. Duttke, S.A. Lacadie, M.M. Ibrahim, C.K. Glass, D.L. Corcoran, C. Benner, S. Heinz, J.T. Kadonaga, U. Ohler, Human promoters are intrinsically directional, Mol. Cell 57 (2015) 674–684.

SUPPLEMENTARY INFORMATION

Core promoter-specific gene regulation: TATA box selectivity and Initiator-dependent bi-directionality of serum response factor-activated transcription

Muyu Xu, Elsie Gonzalez-Hurtado, and Ernest Martinez

Department of Biochemistry, University of California, Riverside, California, 92521, USA

Supplementary Materials and Methods

The sequences of synthetic oligos used for cloning were:

| Oligos	Sequences (5'-3')
UAS^{ACTB}(-120 to -40)	GCGAAGCCGGTGAGTGAGCGGCGCGGGGCCAATCAGCGTG CGCCGTTCCGAAAGTTGCCTTTTATGGCTCGAGCGGCCGC
Mut.CCAAT Forward	CGCGAAGCCGGTGAGTGAGCGGCGCGGGGGATATCAGCGT GCGCCGTTCCGAAAGTTGCCTTTTATGGCTCGAGCGGCCGC GGTAC
Mut.CCAAT Reverse	CGCGGCCGCTCGAGCCATAAAAGGCAACTTTCGGAACGGC GCACGCTGATATCCCCCGCGCCGCTCACTCACCGGCTTCGC GGTAC
Mut.CArG Forward	CGCGAAGCCGGTGAGTGAGCGGCGCGGGGCCAATCAGCGT GCGCCGTTCCGAAAGTTGCCTTTTAGATCTCGAGCGGCCGC GGTAC
Mut.CArG Reverse	CGCGGCCGCTCGAGATCTAAAAGGCAACTTTCGGAACGGCG CACGCTGATTGGCCCCGCGCCGCTCACTCACCGGCTTCGCG GTAC
Mut.Both Forward	CGCGAAGCCGGTGAGTGAGCGGCGCGGGGGATATCAGCGT GCGCCGTTCCGAAAGTTGCCTTTTAGATCTCGAGCGGCCGC GGTAC
Mut.Both Reverse	CGCGGCCGCTCGAGATCTAAAAGGCAACTTTCGGAACGGCG CACGCTGATATCCCCCGCGCCGCTCACTCACCGGCTTCGCG GTAC
5xCArG Forwad	ATCGGGTACCCCTTTTATGGAGCCTTTTATGGAGCCTTTTAT GGAGCCTTTTATGGAGCCTTTTATGGGGTACCATCG
5xCArG Reverse	CGATGGTACCCCATAAAAGGCTCCATAAAAGGCTCCATAAA AGGCTCCATAAAAGGCTCCATAAAAGGGGTACCCGAT
5xCCAAT Forward	ATCGGGTACCGGGCCAATCAGAGGGGCCAATCAGAGGGGC CAATCAGAGGGGCCAATCAGAGGGGCCAATCAGGGTACCA TCG
5xCCAAT Reverse	CGATGGTACCCTGATTGGCCCCTCTGATTGGCCCCTCTGATT GGCCCCTCTGATTGGCCCCTCTGATTGGCCCGGTACCCGAT
5xGC Forward	ATCGGGTACCGGGGCGGGGCAGGGGGCGGGGCAGGGGGCG GGGCAGGGGGCGGGGCAGGGGGCGGGGCGGTACCATCG
5xGC Reverse	CGATGGTACCGCCCCGCCCCCTGCCCCGCCCCCTGCCCCGC CCCCTGCCCCGCCCCCTGCCCCGCCCCGGTACCCGAT
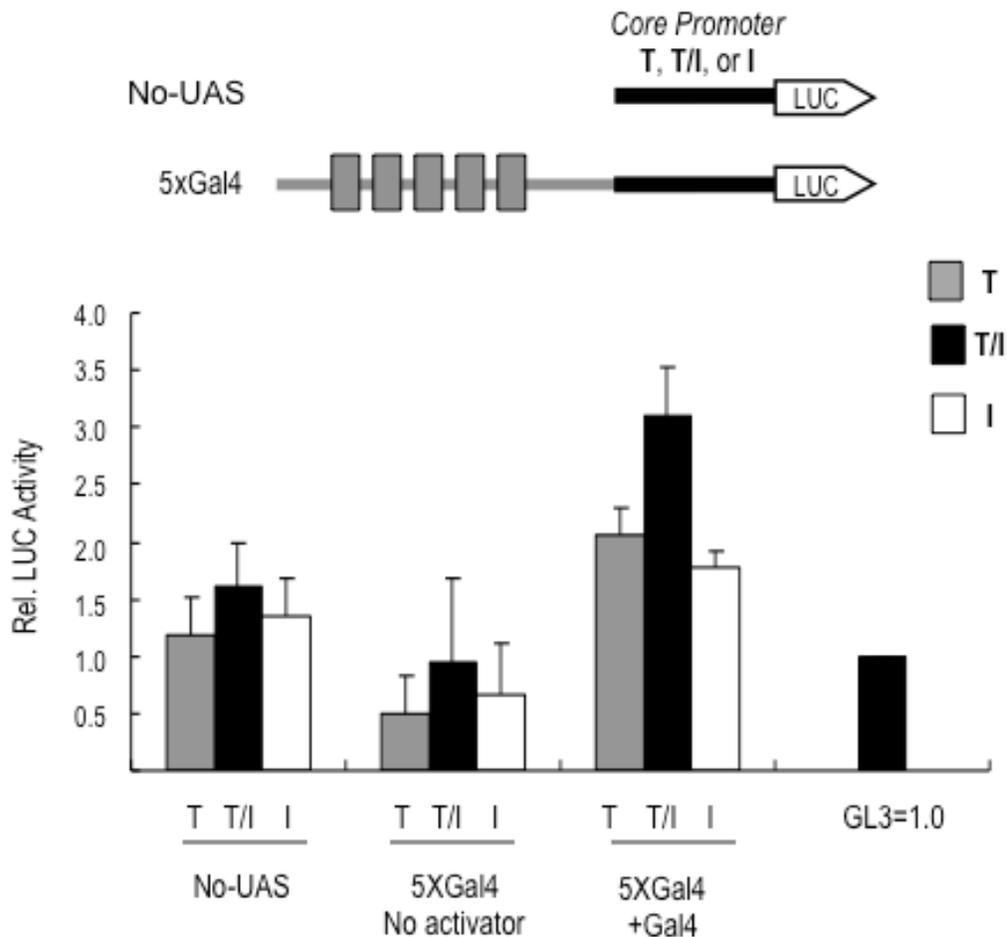
Figure S1 (A): The Gal4 DNA-binding domain (1-147) has a weak activating function that is not core promoter-specific.
HEK293 cells were transfected with luciferase reporter plasmids lacking (No-UAS) or containing 5 Gal4 binding sites (5xGal4) upstream of TATA (T), TATA/INR (T/I) or INR (I) core promoters (schematized in the top diagram) and with or without an expression vector for the yeast Gal4 (1-147) DNA-binding domain. Relative luciferase activities are shown in the bottom part and were normalized to the activity of the internal control pCMV-β-Gal plasmid and to the activity of the promoter-less pGL3 vector (arbitrarily set to 1). Results are means ±S.D. of 3 independent experiments in triplicates.
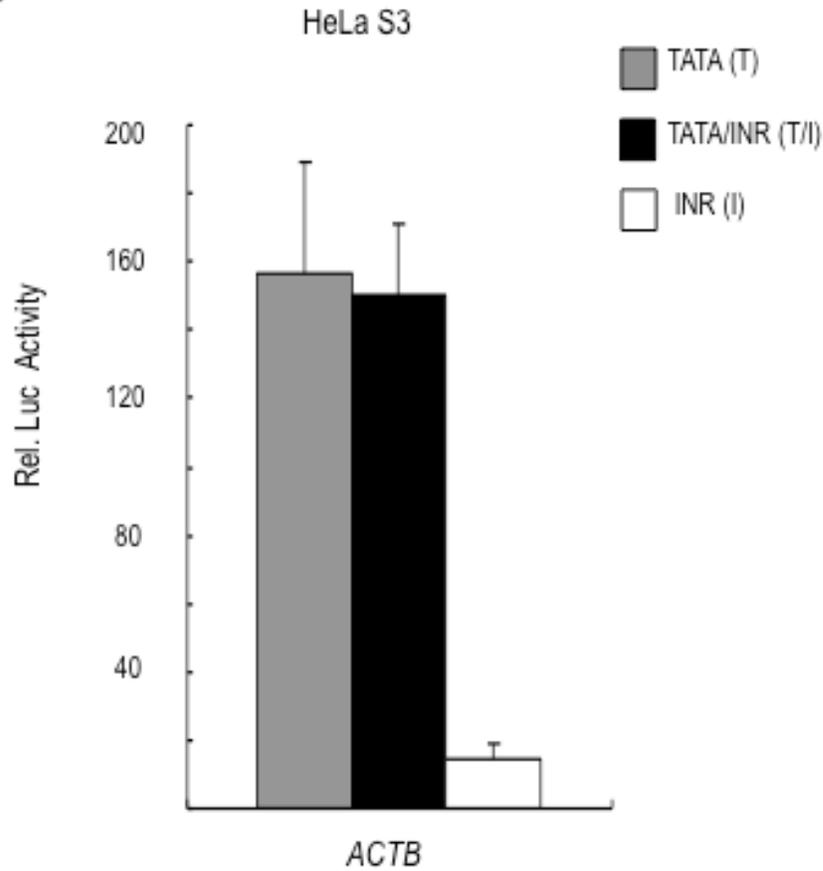
Figure S1 (B): The UAS of *ACTB* is TATA-selective in human HeLa cells
Luciferase reporters containing the UAS of *ACTB* upstream of TATA, TATA/INR or INR heterologous core promoters were transfected into HeLa S3 cells. Relative luciferase activities are shown and were normalized to the β-galactosidase activity of the internal control pCMV-β-Gal plasmid and to the activity of the promoter-less pGL3 vector (arbitrarily set to 1). Results are means ±S.D. of 3 independent experiments in triplicates.
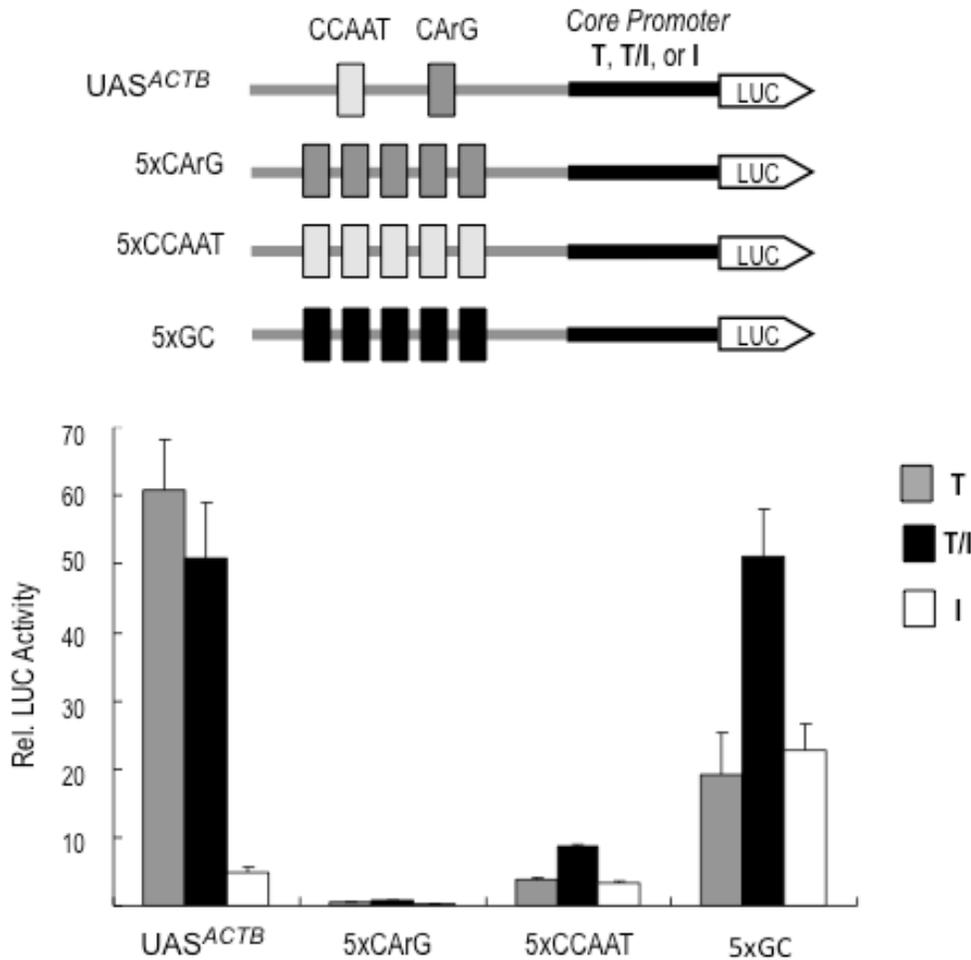
Figure S2: The activity of the native *ACTB* UAS sequence cannot be recapitulated by five copies of either its CArG box or CCAAT box, or by 5 copies of a consensus SP1 binging site. HEK293 cells were transfected with luciferase reporter plasmids containing the wild type *ACTB* UAS (UAS*ACTB*) or five copies of its CCAAT box (5xCCAAT) or CArG-box (5xCArG), or five GC-boxes (5xGC; consensus SP1 sites) upstream of either the TATA, TATA/INR or INR core promoter-luciferase gene (shown in top diagram). Relative luciferase activities are shown in the bottom part and were normalized to the activity of the internal control pCMV-β-Gal plasmid and to the activity of the promoter-less pGL3 vector (arbitrarily set to 1). Results are means ±S.D. of 3 independent experiments in triplicates.

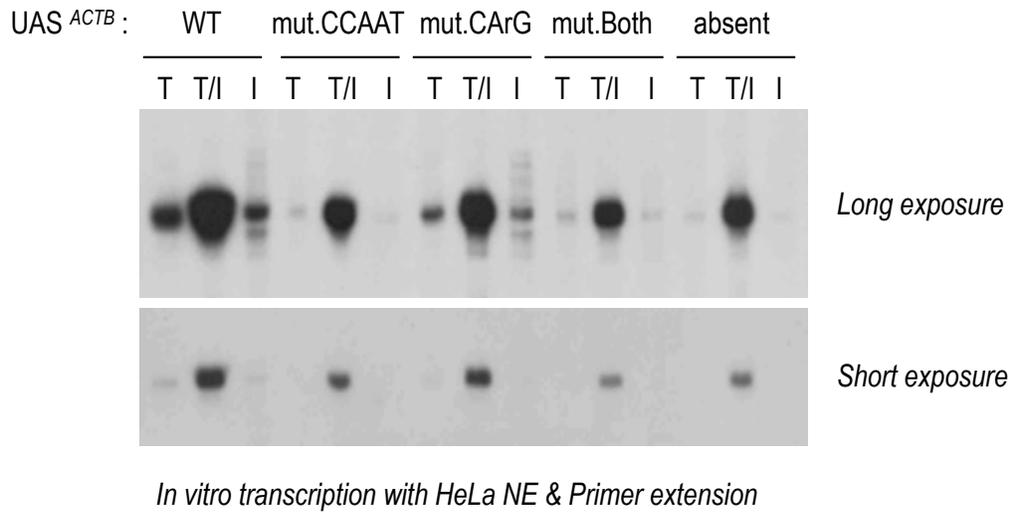*In vitro transcription with HeLa NE & Primer extension*

Figure S3: A factor for TATA-specific activation by the UAS of *ACTB* is limiting in HeLa nuclear extracts.
In vitro transcription reactions were performed using 60 μg HeLa nuclear extract and 40 fmol of the indicated promoter-luciferase plasmids; the promoters are depicted in the top part of Figure 3. Promoters containing either a TATA box only (T), an INR only (I) or both elements (T/I) but without any upstream activating sequence is indicated by "absent". Top and bottom panels are two different exposures of the X-ray film for same in vitro transcription experiment. Note that activation of the TATA-only (T) promoter relative to the TATA/INR (T/I) promoter by the UAS of *ACTB* is less efficient in nuclear extracts compared to transfected cells (Fig. 3 and Fig.S1B).
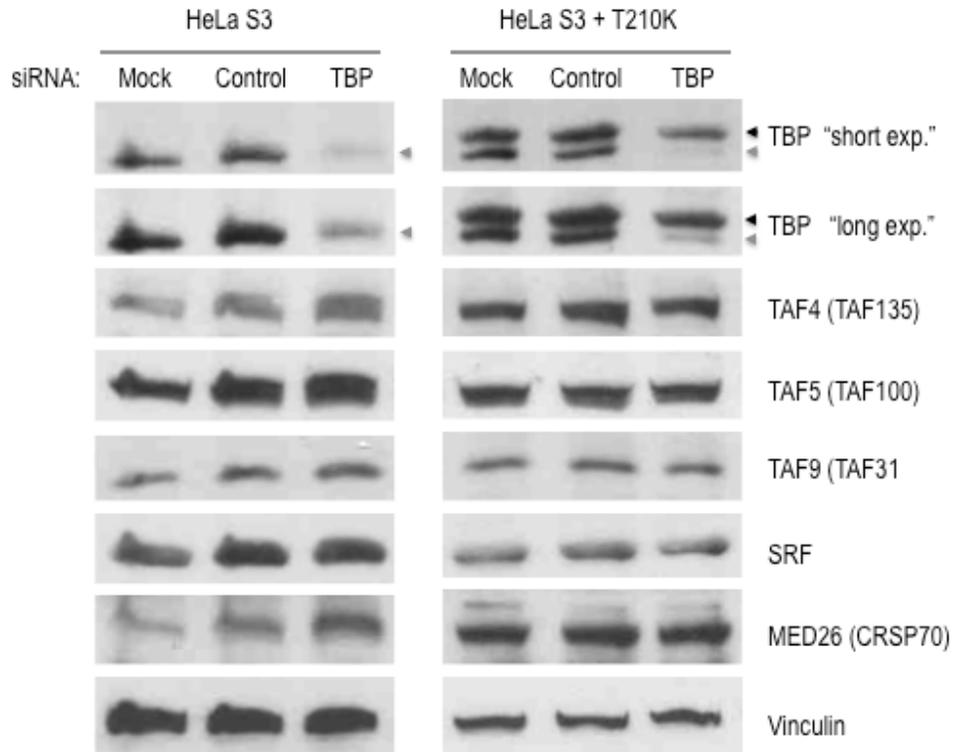
Figure S4: Transient TBP knockdown does not affect expression of TAFs and other cellular proteins.

HeLa S3 or the derivative HeLa S3+T210K cell line were mock transfected or transfected with a Control siRNA or a siRNA targeting TBP 3'UTR, and cellular extracts were analyzed by Western blotting with specific antibodies for the indicated proteins. Ectopic HA-tagged TBP-T210K (closed arrowhead; slow migrating band) and endogenous TBP (gray arrowhead; faster migrating band) are indicated with arrowheads. Two different autoradiographic exposures (short and long exp.) of the same TBP Western blot stripe are shown.