

A core promoter element downstream of the TATA box that is recognized by TFIIB

Wensheng Deng and Stefan G.E. Roberts¹

Faculty of Life Sciences, University of Manchester, Manchester M13 9PT, United Kingdom

We have defined a core promoter element downstream of the TATA box that is recognized by TFIIB. This involves a DNA-binding domain in TFIIB that is distinct from the helix–turn–helix motif (which recognizes an element upstream of the TATA box). The TFIIB recognition element we describe regulates transcription in a manner that is promoter context-dependent, particularly with respect to the TFIIB recognition element that is located upstream of the TATA box. Thus TFIIB can recognize two distinct sequence elements that flank the TATA box, employing independent DNA-binding motifs and cooperating in the regulation of transcription.

Supplemental material is available at <http://www.genesdev.org>.

Received March 6, 2005; revised version accepted August 16, 2005.

The core promoter of genes transcribed by RNA polymerase II (pol II) contains DNA sequence elements that are recognized by the general transcription machinery (Butler and Kadonaga 2002; Smale and Kadonaga 2003; Hahn 2004). These elements help to direct and orient the preinitiation complex (PIC) at the promoter. In addition, they play a critical role in the regulation of transcription (Hochheimer and Tjian 2003; Basehoar et al. 2004; Müller and Tora 2004). The best-characterized core promoter element is the TATA box, which is recognized by the TATA-binding protein (TBP) subunit of TFIID and nucleates PIC formation. However, many promoters do not contain a TATA box, and thus alternative mechanisms are employed to initiate PIC formation. The downstream promoter element (DPE) plays a major role at TATA-less promoters and is recognized by components of TFIID other than TBP. In addition, the initiator element (Inr), which spans the region where mRNA synthesis begins, can nucleate PIC formation via components of TFIID. The Inr can function independently or in combination with a TATA element, exerting a positive effect on transcription. In addition, a recently reported core promoter sequence, the motif-ten-element (MTE), can cooperate with the initiator to stimulate transcription (Lim et al. 2004).

The general transcription factor TFIIB can also engage in sequence-specific contact with the core promoter (Lagrange et al. 1998; Qureshi and Jackson 1998). This

TFIIB-recognition element (BRE) is located immediately upstream of the TATA element in a subset of promoters. A helix–turn–helix (HTH) motif at the C terminus of TFIIB mediates contact with the BRE (Lagrange et al. 1998; Tsai and Sigler 2000). The TFIIB–BRE interaction facilitates the assembly of a TFIIB–TBP–TATA complex (Lagrange et al. 1998; Qureshi and Jackson 1998). In addition, the TFIIB–BRE interaction can orient the PIC on the promoter (Bell et al. 1999; Littlefield et al. 1999; Tsai and Sigler 2000). The BRE acts to increase basal transcription in a system composed of purified factors (Lagrange et al. 1998). In crude nuclear extracts and in living cells, the BRE acts to suppress basal transcription, but has no effect on the level of transcription attained in the presence of an activator protein (Evans et al. 2001; Chen and Manley 2003). Interestingly, the HTH motif is not conserved in yeast or plant TFIIB (Lagrange et al. 1998; Tsai and Sigler 2000).

Previous crystallographic and biochemical studies have indicated that TFIIB also makes sequence-specific DNA contacts downstream of the TATA box (Tsai and Sigler 2000; Fairley et al. 2002; Chen and Manley 2003). Specifically, both the crystal structure of a TFIIB–TBP complex at the Adenovirus major late promoter and also methylation interference analysis of the Adenovirus E4 demonstrate specific contact between TFIIB and two different nucleotides downstream of the TATA element. Here we define a core promoter sequence element downstream of TATA that is contacted by TFIIB. We demonstrate that this contact stabilizes a TFIIB–TBP–promoter complex and modulates promoter strength of a selection of core promoters both *in vitro* using nuclear extracts and in living cells.

Results and Discussion

We have previously provided biochemical evidence that TFIIB contacts bases downstream of the TATA box at the Adenovirus E4 (AdE4) promoter (Fairley et al. 2002). We therefore sought to determine the extent of this contact by derivation and analysis of an optimal DNA sequence. It was first necessary to define the boundaries of the AdE4 TATA element, which was performed by DNase I analysis in combination with permanganate and methylation interference (data not shown).

The TATA box of the AdE4 promoter is shown in bold in Figure 1A. Downstream of TATA, two Gs on the negative strand (indicated by solid circles) and a DNase I hypersensitive region on the positive strand (asterisks) that we previously determined as regions of TFIIB-specific contact are shown (Fairley et al. 2002). Using this information, we generated a library containing randomized bases between positions –23 and +17 within the AdE4 core promoter (bases –51 and +12) (Fig. 1A). This DNA was radiolabeled and then used in bandshift assays with recombinant human TBP and TFIIB. Complexes were isolated, the DNA purified and amplified, and the process repeated to a total of seven rounds of selection. Figure 1B shows a bandshift to analyze TFIIB–TBP–promoter complex formation using either the original randomized pool (Random) or the enriched sequence pool after seven rounds of selection (7th round). The enriched pool exhibited a greater affinity in TFIIB–TBP–promoter complex formation, consistent with the selection of op-

[*Keywords:* Transcription; RNA polymerase II; TFIIB; promoter]

¹Corresponding author.

E-MAIL stefan.roberts@manchester.ac.uk; FAX 0161-275-5082.

Article and publication are at <http://www.genesdev.org/cgi/doi/10.1101/gad.342405>.

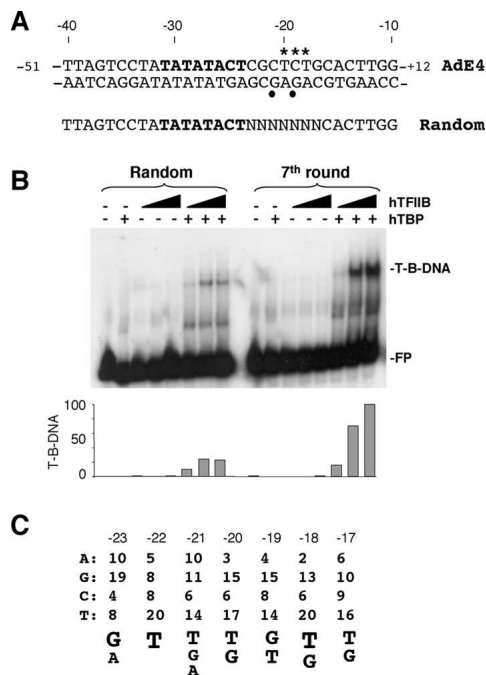


Figure 1. A core promoter element downstream of the TATA box. (A) Sequence of the Adenovirus E4 (Ade4) promoter, both positive and negative strands (nucleotides -51 to +12). The TATA box, as determined experimentally, is shown in bold on the positive strand. The randomized bases (N) downstream of the TATA box (bold) are shown *below*. (B) Bandshift assay using the radiolabeled randomized library (Random) compared with the DNA recovered after seven rounds of selection (7th round). Free probe (FP) and TFIIB-TBP-DNA complex (T-B-DNA) are indicated. Two-hundred nanograms of hTBP and increasing amounts of hTFIIB (4, 20, and 100 ng) were included in the reaction. *Below* the autoradiograph, the data have been quantitated and graphed. (C) Forty-one different clones were sequenced after seven rounds of selection. The frequency of occurrence of each base within the randomized region is shown. *Below*, the consensus sequence derived from the selection is shown with the size of each letter reflecting the frequency (where >9) of the occurrence of the base at each position.

timal sequences. The DNA fragments were then cloned and 41 promoter derivatives were sequenced (Fig. 1C). The selected sequences contain a strong representation of G and T bases and a striking preference against A (especially between bases -17 and -20). Using these data we derived a consensus sequence 5'-G/A-T-T/G/A-T/G-G/T-T/G-T/G-3'. The wild-type Ade4 conforms to this consensus in only three of the seven bases. However, consistent with the selected bases, it does not contain any As within the region -17 to -20. The sequence derived in this study bears no similarity to the BRE, which is immediately upstream of the TATA element in some promoters. From here on we will refer to the BRE identified by Lagrange et al. (1998) as the upstream BRE (BRE^u) and the sequence characterized here as the downstream BRE (BRE^d).

To confirm that the selection experiment had indeed derived a high-affinity BRE^d sequence, we used this information to generate an optimal and a defective BRE^d within the Ade4 promoter (opE4 and dfe4) (Fig. 2A). To facilitate subsequent analysis of the promoter derivatives by methylation interference, the Gs at -19 and -21 (negative strand) were not subject to mutation. Thus the opE4 derivative was generated by mutation of bases -17,

-22, and -23 to the most frequently observed bases derived from the selection (T, T, and G, respectively). The dfe4 derivative was generated by mutation of nucleotides -18 and -20 to A, which was the least favored base at these positions. These promoters were radiolabeled and used in bandshift analysis with hTBP and increasing amounts of hTFIIB (Fig. 2B). Consistent with the selection data, the opE4 derivative showed an enhanced TFIIB-TBP-promoter formation when compared with the wild-type Ade4 promoter. In contrast, the dfe4 derivative showed a poor affinity in TFIIB-TBP-promoter complex formation. Thus, the selection has indeed derived a core promoter sequence preference that can drive formation of a TFIIB-TBP-promoter complex.

We next performed methylation interference to analyze contact at G-19 and G-21 within the opE4 and dfe4 promoter derivatives (negative strand) (Fig. 2C). The data show that methylation of G-19 mediates greater interference at the opE4 derivative when compared with the wild-type Ade4 promoter (wtE4; closed circles). In contrast, and consistent with the selection results and the bandshift analysis, the dfe4 derivative shows severely reduced methylation interference at G-19. We note that methylation interference was not lost at G-21, suggesting that this contact is maintained within the BRE^d mutants that we generated. Permanganate interference was used to ascertain that TBP contact with the opE4 and dfe4 derivatives was not affected by the mutations that we introduced (data not shown).

The wild-type Ade4 promoter was next compared with the opE4 and dfe4 derivatives in their ability to support transcription *in vitro* in a HeLa nuclear extract (Fig. 3A,B). In both the absence (part A) and presence (part B) of an activator the Ade4 derivative that contains an optimized BRE^d (opE4) showed an enhanced level of

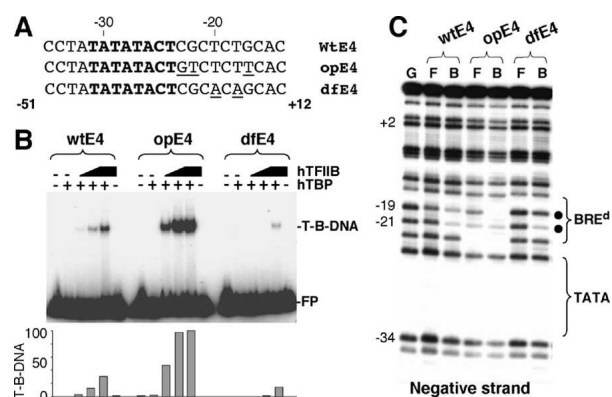


Figure 2. The BRE^d modulates assembly of a TFIIB-TBP-promoter complex. (A) The wild-type Ade4 (wtE4) sequence is shown along with derivatives containing a BRE^d sequence that has either been optimized (opE4) or rendered low affinity (dfe4). (B) Bandshift assay comparing the assembly of a TFIIB-TBP-promoter complex with the wild-type Ade4 promoter (wtE4) and optimized (opE4) or defective (dfe4) derivatives. Free probe (FP) and TFIIB-TBP-DNA complex (T-B-DNA) are indicated. TBP was 50 ng and TFIIB was 1.25, 5, and 20 ng. (C) Methylation interference assay analyzing the wild-type Ade4 promoter (wtE4) and optimized (opE4) or defective (dfe4) derivatives. (G) G-track generated by cleavage of the partially methylated wild-type Ade4 promoter with piperidine; (F) free probe; (B) TFIIB-TBP-bound DNA. Sites of methylation interference are denoted by a solid circle. The region of the TATA box and BRE^d are indicated. The numbers at *left* represent the position relative to the transcription start site.

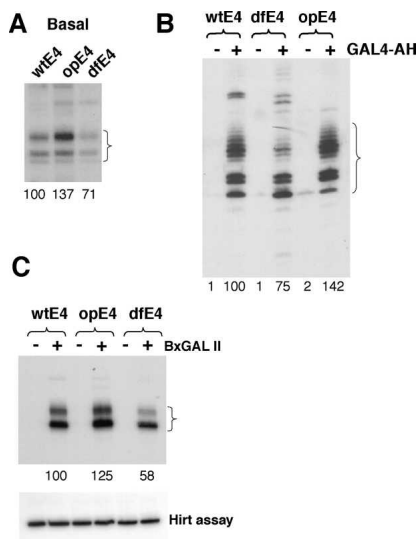


Figure 3. The sequence of the BRE^d determines promoter strength. (A) An in vitro transcription assay using the wild-type AdE4 promoter (wtE4) and optimized (opE4) or defective (dfE4) derivatives. Transcripts were subject to primer extension with a radiolabeled primer, resolved by denaturing electrophoresis, and detected by autoradiography. A bracket indicates the accurately initiated transcripts, and numbers below are quantitation of the transcripts relative to wtE4. (B) As in A, but in the absence or presence of the transcriptional activator GAL4-AH (50 ng). In addition, transcripts were resolved on a sequencing gel. (C) The AdE4 derivatives were tested in a transient transfection assay in the absence and presence of the activator BxGAL II. Transcripts were detected by primer extension of total RNA. A Hirt assay (below) was used to detect (by primer extension) plasmid DNA recovered from the transfected cells.

transcription when compared with the wild-type AdE4 promoter (wtE4), which in turn exhibited a greater level of transcription than that observed with the derivative containing a defective BRE^d (dfE4). In addition, we note that the strength of the BRE^d has significant effects on transcription start site selection (Fig. 3B). Specifically, the transcription start site pattern at the derivative containing a defective BRE^d (dfE4) is more diffuse than that observed with the AdE4 derivative containing an optimal BRE^d (opE4), while the wild-type AdE4 promoter lies in between. Interestingly, our previous studies showed that DNA flanking the TATA box modulates TFIIB conformation in a TFIIB-TBP-promoter complex and that this might affect transcription start site selection (Fairley et al. 2002).

We also tested the AdE4 BRE^d derivatives in living cells (Fig. 3C). The wtE4, opE4, and dfE4 derivatives were transfected into embryonic kidney 293 cells either alone or with a plasmid driving expression of the activator BxGALII. Total RNA was prepared and AdE4 transcripts were analyzed by primer extension. Consistent with our in vitro transcription data, the opE4 derivative showed a greater level of transcription than that observed with the wtE4 promoter and the dfE4 derivative was compromised for transcription. Changes in transcription start site use at the different E4 derivatives were not evident, consistent with our previous report of constrained start site selection at this promoter in vivo (Hawkes and Roberts 1999).

We next turned our attention to the motif in TFIIB that mediates contact with the BRE^d. Deletion analyses

confirmed that the motif was within the core domain of TFIIB (Supplementary Fig. 1A). A previous crystal structure of the TFIIB-TBP-AdML promoter complex demonstrated that a loop in the first direct repeat of TFIIB between helices BH2 and BH3 mediated contact through the minor groove downstream of TATA in the AdML promoter (Fig. 4A; Tsai and Sigler 2000). In this structure Gly153 (G153) mediated base-specific contact and Arg154 (R154) exhibited water-mediated contact within the minor groove. This was further stabilized by Lys152, Arg154, Ala155, and Asn156 through contacts with the DNA backbone. Based on this information we generated a single amino acid substitution mutant of TFIIB (G153Q) and also a double amino acid substitution mutant (G153Q:R154A). These mutant TFIIB derivatives were tested in a bandshift assay to determine their ability to form a TFIIB-TBP-promoter complex (Fig. 4B). Compared with wild-type TFIIB, the G153Q mutant was compromised in its ability to form a TFIIB-TBP-AdE4 complex, and a further reduction was seen with the double substitution TFIIB mutant (G153Q:R154A).

We next reasoned that, if the above mutations affect the ability of TFIIB to interact with the BRE^d, then, in contrast to wild-type TFIIB, the TFIIB mutant G153Q:R154A should show less preference for an optimal over a defective BRE^d sequence. We therefore analyzed TFIIB(G153Q:R154A)-TBP-promoter complex formation at the wild-type AdE4 promoter (wtE4) alongside the opE4 and dfE4 derivatives (Fig. 4C). Significantly, TFIIB G153Q:R154A did not show a reduced ability to assemble at the dfE4 promoter derivative when compared with the wild-type AdE4 promoter (wtE4) or the opE4 derivative. Furthermore, wild-type TFIIB, TFIIB G153Q, and TFIIB G153Q:R154A did not show a significant difference in their ability to form a complex when a BRE^d was not present within the promoter (dfE4) (Fig. 4D). We also performed methylation interference analysis, which revealed that TFIIB G153Q and TFIIB G153Q:R154A both showed reduced interference at both G-19 and G-21 when compared with wild-type TFIIB (Fig. 4E). Taken together, these data demonstrate that the loop between TFIIB helices BH2 and BH3 mediates base-specific contact with the BRE^d.

The TFIIB derivatives G153Q and G153Q:R154A were next tested for their ability to support transcription at the AdE4 promoter. A TFIIB-depleted HeLa nuclear extract was supplemented with either wild-type TFIIB or one of the TFIIB derivatives G153Q and G153Q:R154A and then used in transcription assays with the AdE4 promoter. Figure 4F shows that, when compared with wild-type TFIIB, both TFIIB G153Q and G153Q:R154A were deficient in supporting transcription, particularly in the presence of an activator. These data are consistent with the defect in promoter assembly with TBP that we observed with these derivatives. We note that the defect in transcription that we observed with the TFIIB mutants is more severe than that observed in a standard transcription assay with the AdE4 derivative containing a defective BRE^d (dfE4) (Fig. 3). This can perhaps be explained by the methylation interference data, which showed that contact is maintained with G-21 in the defective AdE4 derivative, whereas both G-21 and G-19 contacts are abated with the TFIIB derivatives containing mutations within the recognition loop. Taken together our results show that, at the AdE4 promoter, the TFIIB-BRE^d inter-

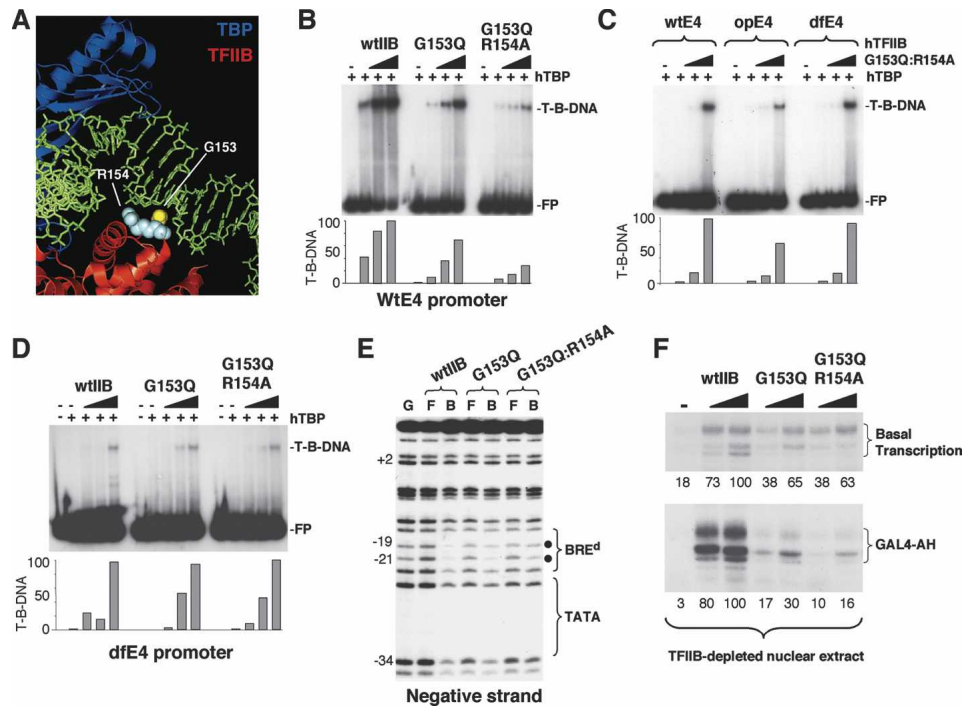


Figure 4. Recognition of the BRE^d by TFIIB. (A) A structural representation of the TFIIB-TBP-AdML promoter complex generated from the coordinates of Tsai and Sigler (2000) using PYMOL. DNA is green, TBP is blue, and TFIIB is red. TFIIB residues G153 (yellow) and R154 (blue) are indicated. (B) Bandshift assay comparing the assembly of wild-type TFIIB or the TFIIB derivatives G153Q and G153Q:R154A with TBP at the wild-type Ade4 promoter. TBP was 50 ng and TFIIB (or derivatives) was 0.4, 2, and 10 ng. Free probe (FP) and TFIIB-TBP-DNA complex (T-B-DNA) are indicated. (C) Bandshift assay comparing the assembly of a TFIIB (G153Q:R154A)-TBP-promoter complex with the wild-type Ade4 promoter (wtE4) and optimized (opE4) or defective (dfE4) derivatives. Amounts of each factor are as in B. (D) Bandshift assay comparing the assembly of wild-type TFIIB or the TFIIB derivatives G153Q and G153Q:R154A with TBP at the dfE4 promoter. Amounts of each factor are as in B. (E) Methylation interference analysis at the wild-type Ade4 promoter was performed with TBP and either wild-type TFIIB or the TFIIB derivatives G153Q and G153Q:R154A. (G) a G-track; (F) free probe; (B) TFIIB (or derivative)-TBP-bound DNA. Sites of methylation interference are denoted by a solid circle. The region of the TATA box and BRE^d are indicated. The numbers at *left* represent the position relative to the transcription start site. (F) TFIIB-depleted HeLa nuclear extract supplemented with either wild-type TFIIB (wtIIB), TFIIB G153Q, or TFIIB G153Q:R154A (10 and 100 ng). The wild-type Ade4 promoter was used as a template and transcripts were detected by primer extension. Numbers *below* each panel are quantitation of the transcripts relative to the maximal level observed with wild-type TFIIB. The *top* panel shows activator-independent basal transcription. The *lower* panel is the same, but includes the transcriptional activator GAL4-AH (50 ng). A Coomassie-stained gel of the wild-type TFIIB and derivative proteins along with an immunoblot showing the depletion of TFIIB from a HeLa nuclear extract are shown in Supplementary Figure 1.

action has a positive effect on both basal and activator-dependent transcription in both nuclear extracts and in living cells. Furthermore, this sequence-specific TFIIB-DNA contact is mediated by a recognition loop in the first direct repeat of TFIIB between helices BH2 and BH3.

To determine the potential prevalence of the BRE^d, we analyzed 140 promoters from the eukaryotic promoter database (Périer et al. 2000). Only promoters that contain a single consensus TATA sequence were chosen to allow accurate location of the BRE^d. Six percent of these core promoters contains at least six out of seven bases of the consensus sequence, 18% contain at least five of seven bases and 37% contain at least four of seven bases (random chance figures would be 1.2%, 2.3%, and 4.6%, respectively). The prevalence of the BRE^d is thus similar to that observed for the BRE^u (Lagrange et al. 1998). In TATA-containing promoters, the BRE^d shows a similar prevalence regardless of the presence of other elements, including the BRE^u. Our selection data also showed that A was significantly selected against between positions -17 and -20. We note that of the 140 eukaryotic promoters analyzed above, at each of the positions between -17 and -20, A occurred at rates of only 14%, 14%, 14%, and 13%. In addition, at promoters with low

nucleotide match to the BRE^d, A is not overrepresented when compared with C, suggesting that As are unlikely to be employed as a negative-acting sequence in place of the BRE^d.

Using the information above we selected the hTPT1 and hCompF core promoters, which contain seven/seven and six/seven base matches to the BRE^d consensus, respectively, to test the functional relevance of the BRE^d to a cellular gene (Fig. 5A,B; see Supplementary Fig. 2 for the promoter sequences). Mutant derivatives of both promoters were generated to render the BRE^d sequence non-functional (df) and the transcriptional activities compared in a HeLa nuclear extract in both the absence and presence of the activator GAL4-AH. Mutation of the BRE^d at both the hTPT1 and hCompF core promoters caused a significant reduction in transcription, confirming a role for the BRE^d in the regulation of cellular genes. We note that both the hTPT1 and hCompF promoters lack an Inr sequence and hence use multiple transcription start sites. We also selected a human promoter (H2AFY) that contains a poor TATA element, but both a consensus Inr and DPE in addition to a six/seven match BRE^d. Mutation of the BRE^d within the H2AFY promoter caused a small, but reproducible, reduction in transcrip-

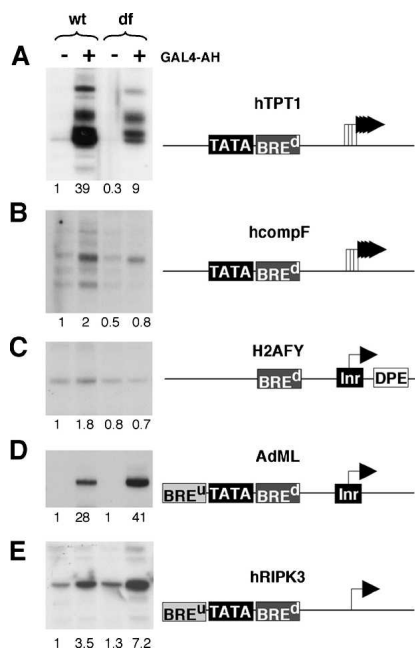


Figure 5. The function of the BRE^d is context-dependent. The core promoters indicated in A–E (wild type; wt) were tested in transcription assays alongside a derivative of each promoter that contained a mutated BRE^d sequence (df). The numbers *below* each panel are quantitation of the level of transcription relative to that observed at basal level for each wild-type promoter. Schematics of each promoter are shown at the *right* with the core promoter elements indicated; arrows depict the transcription start sites. The core promoters were compared in an *in vitro* transcription assay with HeLa nuclear extract in the absence or presence of the activator GAL4-AH (50 ng). Full details for each core promoter and specific mutations are shown in Supplementary Figure 2.

tion, although we also note that the promoter elicited a poor response to GAL4-AH (Fig. 5C).

The transcriptional effects of the BRE^d are distinct from those observed with the BRE^u in nuclear extracts and living cells (Evans et al. 2001; Chen and Manley 2003). While the BRE^u has a negative effect on basal transcription of the AdML promoter in nuclear extracts, it does not affect activator-dependent transcription at this promoter under the same conditions. In contrast, our data so far have shown, in the same transcription system, that the BRE^d has a positive effect at the AdE4, hTPT1, hCompF, and H2AFY promoters.

The Adenovirus major late (AdML) promoter contains both a BRE^u and a six/seven match to the consensus BRE^d. A recent study by Chen and Manley (2003) analyzed a mutant AdML derivative that, based on our selection studies, we predict would weaken the BRE^d. They found that this mutant enhanced activator-dependent transcription. We therefore generated an AdML derivative that disrupts the BRE^d in the same way as the other promoters analyzed above. Wild-type AdML and the AdML derivative containing a defective BRE^d were tested in an *in vitro* transcription assay in the absence and presence of the activator GAL4-AH (Fig. 5D). In agreement with the study of Chen and Manley (2003), disruption of the AdML BRE^d enhanced transcriptional activation. We obtained comparable data in a transient transfection assay with the wtML and dfML derivatives (Supplementary Fig. 3). We also analyzed the effect of

BRE^d mutation on another BRE^u-containing promoter (hRIPK3) (Fig. 5E). As we observed with the AdML, mutation of the BRE^d element in the hRIPK3 promoter caused an increase in transcriptional activation. Thus, in a promoter context-dependent manner, the BRE^d can elicit either a positive or negative effect on transcription. It is therefore likely that there is cross-talk between the TFIIB recognition elements via their contacts with TFIIB.

Considering that both BREs can act to stabilize the assembly of TFIIB at the promoter, it is intriguing that they can also elicit a negative effect on transcription at some promoters when analyzed in nuclear extracts and living cells. This could involve a role for inhibitory transcription factors such as Dr1/NC2 and Mot1 (for review, see Smale and Kadonaga 2003). In addition, while GTF-core promoter interactions can enhance PIC formation, they might also impede the transition from initiation to promoter escape. Our current work therefore adds more weight to the notion that it is the blend of core promoter elements rather than individual elements per se that direct the final transcriptional outcome.

Interestingly, the TFIIB HTH motif, which contacts the BRE^u, is not conserved among all eukaryotes (it is absent in yeast and plants). However, the TFIIB recognition loop containing G153 and R154 is conserved in all eukaryotes (Tsai and Sigler 2000). Indeed, at the AdE4 promoter, yeast TFIIB shows the same methylation interference pattern as that observed with human TFIIB (see Supplementary Fig. 4). Based on the studies of Tsai and Sigler (2000), the TFIIB–DNA contact with the BRE^d takes place via the minor groove, while that with the BRE^u takes place through the major groove. It is interesting that TFIIB has evolved to mediate two independent DNA sequence-specific contacts with the promoter. This is reminiscent of prokaryotic σ factors, which also contain two DNA-binding motifs that contact the –35 and –10 elements within bacterial promoters (for review, see Borukhov and Severinov 2002). Further characterization of the eukaryotic TFIIB recognition elements will determine their precise relationship and the extent of their use in core promoter function *in vivo*.

Materials and methods

Plasmids and proteins

G9E4 containing nucleotides –51 to +12 from Adenovirus E4 promoter has been described previously (Fairley et al. 2002). The expression vectors pEthTFIIB, pETHIS-hTBP, and pETGal4AH have been described before (Lin et al. 1988; Hawkes and Roberts 1999; Fairley et al. 2002). Recombinant TFIIB (and derivatives) and GAL4-AH were purified as previously described (Ha et al. 1993; Lin et al. 1988). Recombinant hTBP was purified with Nickel agarose as described by the manufacturer (Qiagen). Promoter and TFIIB derivatives were generated by using the Quickchange site-directed mutagenesis kit as described by the manufacturer (Stratagene).

Random selection and bandshift assays

The oligonucleotides used for random selection were synthesized as follows: AdE4 DNA template, 5'-TACGTCATTTTTAGTCCTATATATA CTNNNNNNNCACTTGGCCCTTTTTTACACTGTGACTG-3'; RSP1, 5'-GGCGGATCCTACGTCATTTTTTAGTCC-3'; RSP2, 5'-GCCGAA-TTCCAGTCACAGTGTA AAAAAGG-3'. This AdE4 DNA template contains seven consecutive randomized nucleotides downstream of the TATA box. PCR primers RSP1 and RSP2 contain restriction sites recognized by BamHI and EcoRI to assist cloning. For the first cycle, 600 ng AdE4 template and 200 ng RSP2 were used to perform primer extension in a 50- μ L reaction mixture containing 50 μ M dGTP, dTTP, dCTP, 25

μM dATP, 1 μL of $\alpha^{32}\text{P}$ dATP (10 μCi) and 5 U of Taq DNA polymerase (Roche). The products of primer extension were purified on a 6% native polyacrylamide gel. The radiolabeled DNA fragments were used in a bandshift assay with recombinant hTFIIB and hTBP as described previously (Maldonado et al. 1990). The band containing TBP-TFIIB complexes was excised after detection with autoradiography, then DNA fragments were recovered by incubating the gel slice with 600 μL ddH₂O overnight at 37°C and precipitation with ethanol. The recovered DNA fragments were used as a PCR template for the next cycle of selection. After a total of seven rounds of selection, the DNA fragments were digested with BamHI and EcoRI, followed by cloning into pGEM3 downstream of nine GAL4 DNA-binding sites.

Methylation interference and transcription assays

Methylation interference analysis was performed as described previously (Ausubel et al. 1995; Evans et al. 2001; Fairley et al. 2002). Transcription assays, transient transfection, Hirt assay, and TFIIB depletion of nuclear extracts were performed as described before (Lin et al. 1988; Hawkes and Roberts 1999; Evans et al. 2001). HeLa cell nuclear extracts were purchased from Computer Cell Culture Centre.

Acknowledgments

We thank Richard Reece, Andy Sharrocks, Paul Shore, and members of the laboratory for comments on the manuscript. This work was funded by the Wellcome Trust (061207/Z/00/Z/CH/TG/dr). S.G.E.R. is a Wellcome Trust Senior Research Fellow.

References

- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. 1995. *Short protocols in molecular biology*, pp. 12.9–12.11. Wiley, New York.
- Basehoar, A.D., Zanton, S.J., and Pugh, B.F. 2004. Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **116**: 699–709.
- Bell S.D., Kosa, P.L., Sigler, P.B., and Jackson, S.P. 1999. Orientation of the transcription preinitiation complex in Archaea. *Proc. Natl. Acad. Sci.* **96**: 13662–13667.
- Borukhov, S. and Severinov, K. 2002. Role of the RNA polymerase σ subunit in transcription initiation. *Res. Microbiol.* **153**: 557–562.
- Butler, J.E. and Kadonaga, J.T. 2002. The RNA polymerase II core promoter: A key component in the regulation of gene expression. *Genes & Dev.* **16**: 2583–2592.
- Chen, Z. and Manley, L. 2003. Core promoter elements and TAFs contribute to the diversity of transcriptional activation in vertebrates. *Mol. Cell. Biol.* **25**: 7350–7362.
- Evans, R., Fairley, J.A., and Roberts, S.G.E. 2001. Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB. *Genes & Dev.* **15**: 2945–2949.
- Fairley, J.A., Evans, R., Hawkes, N.A., and Roberts, S.G.E. 2002. Core promoter-dependent TFIIB conformation and a role for TFIIB conformation in transcription start site selection. *Mol. Cell. Biol.* **22**: 6697–6705.
- Ha, I., Roberts, S.G.E., Maldonado, E., Sun, X., Kim, L.-U., Green, M.R., and Reinberg, D. 1993. Multiple functional domains of general transcription factor IIB: Distinct interactions with two general transcription factors and RNA Polymerase II. *Genes & Dev.* **7**: 1021–1032.
- Hahn, S. 2004. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* **11**: 394–403.
- Hawkes, N.A. and Roberts, S.G.E. 1999. The role of human TFIIB in transcription start site selection in vitro and in vivo. *J. Biol. Chem.* **274**: 14337–14343.
- Hochheimer, A. and Tjian, R. 2003. Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. *Genes & Dev.* **17**: 1309–1320.
- Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., and Ebright, R.H. 1998. New core promoter element in RNA polymerase II-dependent transcription: Sequence specific DNA binding by transcription factor IIB. *Genes & Dev.* **12**: 34–44.
- Lim, C.Y., Santoso, B., Boulay, T., Dong, E., Ohler, U. and Kadonaga, J.T. 2004. The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes & Dev.* **18**: 1606–1617.
- Lin, Y.-S., Carey, M., Ptashne, M., and Green, M.R. 1988. GAL4 derivatives function alone and synergistically with mammalian activators *in vitro*. *Cell* **54**: 659–664.
- Littlefield, O., Korkhin, Y., and Sigler, P.B. 1999. The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proc. Natl. Acad. Sci.* **96**: 13668–13673.
- Maldonado, E., Ha, I., Cortes, P., Weis, L., and Reinberg, D. 1990. Role of transcription factors IIA, IID and IIB during formation of a transcription competent complex. *Mol. Cell. Biol.* **10**: 6335–6347.
- Müller, F. and Tora, L. 2004. The multicoloured world of promoter recognition complexes. *EMBO J.* **23**: 2–8.
- Périer, R.C., Praz, V., Junier, T., Bonnard, C., and Bucher, P. 2000. The Eukaryotic Promoter Database (EPD). *Nucleic Acids Res.* **28**: 302–303.
- Qureshi, S.A. and Jackson, S.P. 1998. Sequence-specific DNA binding by the *S. shibatae* TFIIB homolog, TFB, and its effect on promoter strength. *Mol. Cell* **1**: 389–400.
- Smale, S.T. and Kadonaga, J.T. 2003. The RNA polymerase II core promoter. *Annu. Rev. Biochem.* **72**: 449–479.
- Tsai, F.T.F. and Sigler, P.B. 2000. Structural basis of preinitiation complex assembly on human Pol II promoters. *EMBO J.* **19**: 25–36.