Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters

Thomas W. Burke and James T. Kadonaga[1]

Department of Biology and Center for Molecular Genetics, University of California at San Diego, La Jolla, California 92093-0347 USA

We describe the identification and characterization of a conserved downstream basal promoter element that is present in a subset of *Drosophila* TATA-box-deficient (TATA-less) promoters by using purified, epitope-tagged TFIID complex (eTFIID) from embryos of transgenic *Drosophila*. DNase I footprinting of the binding of eTFIID to TATA-less promoters revealed that the factor protected a region that extended from the initiation site sequence (about +1) to ~35 nucleotides downstream of the RNA start site. In contrast, there was no apparent upstream DNase I protection or hypersensitivity induced by eTFIID in the −25 to −30 region at which TATA motifs are typically located. Further studies revealed a conserved sequence motif, $^A/_GG^A/_TCGTG$, termed the downstream promoter element (DPE), which is located ~30 nucleotides downstream of the RNA start site of many TATA-less promoters. DNase I footprinting and in vitro transcription experiments revealed that a DPE in its normal downstream location is necessary for transcription of DPE-containing TATA-less promoters and can compensate for the disruption of an upstream TATA box of a TATA-containing promoter. Moreover, a systematic mutational analysis of DNA sequences that encompass the DPE confirmed the importance of the consensus DPE sequence motif for basal transcription and further supports the postulate that the DPE is a distinct, downstream basal promoter element. These results suggest that the DPE acts in conjunction with the initiation site sequence to provide a binding site for TFIID in the absence of a TATA box to mediate transcription of TATA-less promoters.

The regulation of gene expression is often controlled at the level of transcription, and thus, the elucidation of the factors and mechanisms involved in basal transcription by RNA polymerase II is of fundamental biological significance (for review, see Conaway and Conaway 1993; Buratowski 1994; Maldonado and Reinberg 1995; Zawel and Reinberg 1995). The current data on basal transcription factors indicate that RNA polymerase II, along with auxiliary factors termed TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH, are required for transcription initiation, whereas the polymerase and TFIIF, TFIIS, and elongin/SIII are involved in transcriptional elongation. It has also been proposed that some of the RNA polymerase II exists as a component of a large complex termed the "holoenzyme" that contains RNA polymerase II (also referred to as the "core" polymerase) and many other factors, which appear to include TFIIB, TFIIE, TFIIF, and TFIIH (see, e.g., Koleske and Young 1995).

In addition to the identification of the basal factors, it is important to study the mechanism of the transcription process. There is, at present, considerable evidence that there are variations in the mechanisms by which basal transcription occurs, and much, if not virtually all, of the information that dictates the specific characteristics of the basal transcription process at a gene appears to be provided by the core promoter sequence. Accordingly, the critical core promoter elements that direct basal transcription have been studied thoroughly. Two well-characterized basal promoter elements are the TATA box and the initiator (Inr) motif. The TATA box is an A/T-rich sequence that is present in some, but not all, promoters at ~25–30 nucleotides upstream of the transcription start site. The TATA element is a binding site for the TATA-box-binding polypeptide (TBP) component of the multisubunit TFIID complex (for review, see Pugh and Tjian 1992; Smale 1994). The consensus sequence for the TATA box is often taken to be TATAAA, but it has been shown that there can be considerable variation in the TATA box sequence (see, e.g., Singer et al. 1990; Wiley et al. 1992; Zenzie-Gregory et al. 1993).

The Inr element was identified as a sequence that encompasses the RNA start site that is sufficient to direct

[1]Corresponding author.

accurate transcription in the absence of a TATA box (see, e.g., Smale and Baltimore 1989; Weis and Reinberg 1992; Smale 1994). The consensus for the Inr element is Py-Py-A_{+1}-N-$^T/_A$-Py-Py (where A_{+1} is the transcription start site) for mammalian genes (Smale and Baltimore 1989; Javahery et al. 1994) and T-C-A_{+1}-$^G/_T$-T-$^T/_C$ for *Drosophila* genes (Hultmark et al. 1986; Purnell et al. 1994; Arkhipova 1995). As with the TATA box element, sequences that resemble the Inr consensus are found in some but not all promoters. Moreover, in addition to the TATA box and Inr core promoter elements, sequences that are immediately downstream (up to about +40) of the RNA start site have been found, in some instances, to be required for efficient basal transcription.

To further our analysis of basal transcription by RNA polymerase II, we have devised an effective means for the purification of epitope-tagged TFIID complex (eTFIID) from transgenic *Drosophila*. TFIID from *Drosophila* (dTFIID) is believed to consist of TBP and at least eight TBP-associated factors (TAFs) (see, e.g., Dynlacht et al. 1991; Hoey et al. 1993; Kokubo et al. 1993a; Chen et al. 1994). Because the binding of TFIID appears to be the first step in the pathway leading to the assembly of the transcription preinitiation complex (Buratowski et al. 1989; Maldonado et al. 1990), we characterized the binding of the purified eTFIID to various TATA-box-containing promoters as well as *Drosophila* TATA-box-deficient (TATA-less) promoters. These binding studies, along with corresponding transcriptional experiments, have led to the identification of a conserved downstream promoter element that is present in many *Drosophila* TATA-less promoters. This new basal promoter element, the DPE (downstream promoter element), appears to act in conjunction with initiation site sequences to provide a binding site for TFIID in the absence of a TATA box to mediate transcription of TATA-less promoters.

Results

Purification of epitope-tagged TFIID from transgenic Drosophila embryos

To obtain purified dTFIID complex for the biochemical analysis of transcription by RNA polymerase II, we have expressed a modified dTBP cDNA that encodes dTBP with an amino-terminal epitope tag, FLAG (Hopp et al. 1988), in *Drosophila*. This epitope-tagging strategy has been used previously to purify the TFIID complex from humans (Zhou et al. 1992; Chiang et al. 1993) and *Saccharomyces cerevisiae* (Poon et al. 1995). For our studies with *Drosophila*, we stably integrated the cDNA encoding the epitope-tagged dTBP (eTBP) into the germ line by using P-element-mediated transformation (Rubin and Spradling 1982). Then, by standard genetic techniques, we obtained homozygous transformed flies that stably and constitutively expressed eTBP. By Western blot analysis with antibodies directed against dTBP, the native TBP to eTBP ratio was estimated to be 3:1 in embryos of these transgenic flies (data not shown). Hence, there was

an excess of native TBP and, presumably, TAFs, relative to eTBP.

We then purified eTFIID from embryos of the transgenic *Drosophila* as outlined in Figure 1A. A nuclear extract was fractionated by P11 phosphocellulose chromatography to separate eTFIID from other eTBP-containing species (see, e.g., Zawel and Reinberg 1995). The eTFIID-containing P11 fraction was then subjected to
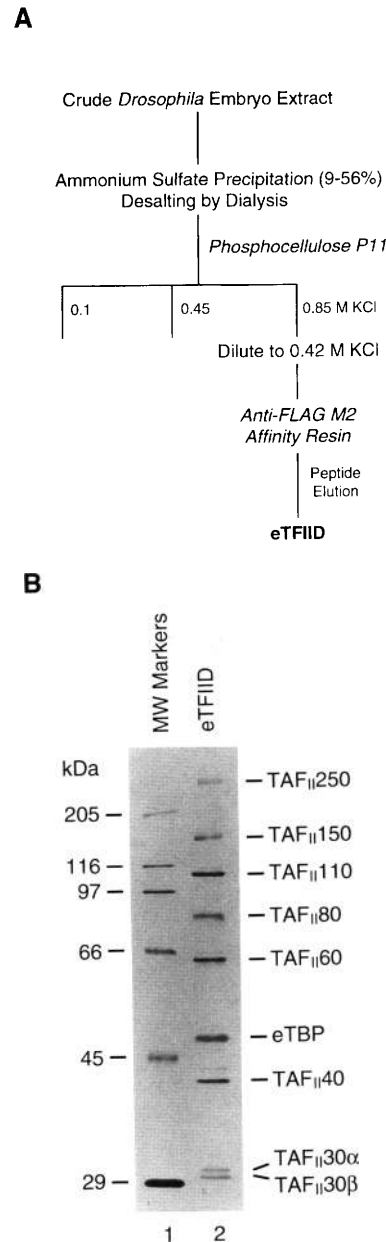

Figure 1. Purification of eTFIID from embryos of transgenic *Drosophila*. (*A*) Scheme for the fractionation and purification of *Drosophila* eTFIID. (*B*) Analysis of eTFIID by 9% polyacrylamide–SDS gel electrophoresis and silver staining. (Lane *1*) Molecular mass standards; (lane *2*) purified eTFIID (500 ng). The identity of each of the subunits of *Drosophila* eTFIID is indicated. TAFs that were smaller than 30 kD were not observed (data not shown).

immunoaffinity purification under nondenaturing conditions with an anti-FLAG M2 affinity resin (IBI/Kodak). The protein composition of the purified eTFIID, as visualized by polyacrylamide–SDS gel electrophoresis and silver staining (Fig. 1B), is identical to that reported for native dTFIID by Tjian and co-workers (see, e.g., Dynlacht et al. 1991; Hoey et al. 1993; Chen et al. 1994) and similar to that described by Nakatani and co-workers (see, e.g., Kokubo et al. 1993a–c; Kokubo et al. 1994). By Western blot analysis, dTFIIA, which is also known to interact with TBP, was not present in our preparations of eTFIID (data not shown). The eTFIID complex was estimated to be ~90% homogeneous, and a summary of the purification steps is given in Table 1. This purification procedure was found to be consistent and reproducible through >20 independent preparations of eTFIID. In all subsequent experiments, the eTFIID was of comparable purity to that shown in Figure 1B.

eTFIID can function in both basal and activated transcription

We first tested whether eTFIID was biochemically active in reconstituted transcription reactions. To analyze basal transcription, we carried out reactions with purified transcription factors (Tyree et al. 1993) and found that eTFIID possessed activity that was comparable to that of dTBP for basal transcription with the adenovirus major late promoter (data not shown). We then investigated the ability of the model activator GAL4–VP16 (Sadowski et al. 1988; Chasman et al. 1989) to activate transcription in reactions reconstituted with eTFIID. In this experiment, reconstituted transcription reactions were performed with either dTBP or eTFIID in the presence or absence of the GAL4–VP16 activator protein with DNA templates that contained either five or zero GAL4 binding sites upstream of the TATA box of the adenovirus E4 (AdE4) promoter (Fig. 2). In the absence of the activator, we observed a lower level of basal transcription with dTBP than with eTFIID under otherwise identical conditions. With dTBP, activation by GAL4–VP16 was not observed, whereas with eTFIID, a mild (2.8-fold) activa-
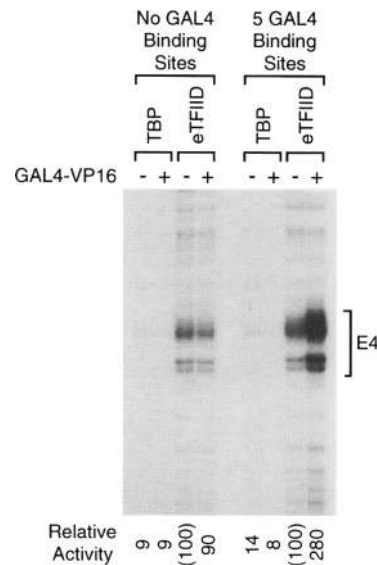

Figure 2. Purified *Drosophila* eTFIID can function for transcription in vitro. Transcription from the AdE4 promoter in the presence or the absence of the GAL4–VP16 activator. Transcription reactions were performed with either pG$_0$E4T (no GAL4 sites) or pG$_5$E4T (five GAL4 sites) template DNAs (Lin et al. 1988; 100 ng) with purified recombinant basal factors [dTFIIA (10 ng), dTFIIB (5 ng), hTFIIE34 (2 ng), hTFIIE56 (10 ng), hTFIIF30 (6 ng), hTFIIF74 (10 ng)], purified calf thymus RNA polymerase II (40 ng), and a partially purified dTFIIH fraction (4 ng). Where noted, dTBP (10 ng), eTFIID (60 ng TFIID complex), and GAL4–VP16 (30 ng) were included in the reactions. The reverse transcription products are indicated by a bracket. The relative transcriptional activity observed for each template DNA is shown below the autoradiographs.

tion of transcription was seen. This amount of transcriptional activation by GAL4–VP16 is comparable to that which we and others have observed with the same promoter construction in the absence of histones or other transcriptional repressors (see, e.g., Pazin et al. 1994). Moreover, as a control, transcriptional activation by GAL4–VP16 with eTFIID was not observed with a DNA

Table 1. *Fractionation and immunopurification of eTFIID from transgenic* Drosophila *embryos*

Fraction	Total volume (ml)	Total protein (mg)	eTBP Western (units)[a]	Relative activity (U/mg)[b]	Yield (%)[b]	Purification (-fold)[b]
Crude nuclear extract	10	257[c]	5670	22	100	1
0.45–0.85 M P11	9	9.2[c]	419	46	7.4	2.1
Anti-FLAG immunopurification	0.1	0.006[d]	115	19,100	2	870

Fractionation and immunopurification were based on 80 grams of embryos.

[a]One Western eTBP unit is an arbitrary value obtained by quantitation of the radioactive bands (by using a Fuji PhosphorImager) in an anti-TBP Western blot assay, as described in Materials and methods.

[b]Because the eTBP that was incorporated into TFIID complexes was a fraction of the total eTBP in the crude extract, the values given for eTBP Western and Relative activity are overestimates of eTFIID in the starting material (crude extract). Hence, the values for Relative activity, Yield, and Fold purification are underestimates of the actual values.

[c]Protein concentration was determined by Coomassie protein assay (Pierce).

[d]Protein concentration was estimated from an SDS–polyacrylamide gel.

template that did not contain GAL4 binding sites (Fig. 2). It thus appears that eTFIID possesses a greater ability than dTBP to function with GAL4–VP16 in the activation of transcription, and hence, the transcriptional properties of eTFIID are similar to those described previously for the native dTFIID complex.

eTFIID binds to both TATA-box-containing promoters and TATA-less promoters

To investigate potential differences in the binding of eTFIID to TATA-containing versus TATA-less promoters, we proceeded to study the binding of eTFIID to different promoters. First, we tested the adenovirus major late (AdML), *Drosophila Krüppel* (*Kr*), and AdE4 promoters, each of which contains a TATA box element. With the AdML promoter (Fig. 3A, left), eTFIID binds to an extended region from the TATA box to ~34 nucleotides downstream of the RNA start site, as seen previously with either partially purified TFIID (see, e.g., Sawadogo and Roeder 1985; Nakajima et al. 1988) or purified TFIID from humans or *Drosophila* (see, e.g., Zhou et al. 1992; Chiang et al. 1993; Verrijzer et al. 1994). This extended TFIID footprint with the AdML promoter is distinct from the footprint observed with purified TBP, which binds to a localized region that encompasses the TATA box but does not extend downstream of the transcription start site (see, e.g., Buratowski et al. 1988; Cavallini et al. 1989; Schmidt et al. 1989; Hoey et al. 1990; Peterson et

al. 1990). The binding of dTBP relative to eTFIID with the *Krüppel* promoter was similar to that seen with the AdML promoter (Fig. 3A, middle). On the other hand, with the AdE4 promoter, the eTFIID footprint did not extend downstream of the TATA box and thus resembled the TBP footprint (Fig. 3A, right). These results are in agreement with the known characteristics of TFIID binding to TATA-containing promoters.

We then sought to analyze further the binding of eTFIID to TATA-less promoters. To this end, we examined two *Drosophila* promoters: the downstream P2 promoter of the *Antennapedia* (*Antp*) gene and the internal promoter of the *jockey* (*joc*) mobile element (Fig. 3B). Previous studies with these TATA-less promoters had suggested that efficient transcription from these promoters did not require sequences upstream of the RNA start site in the location where a TATA box would normally be located (Mizrokhi et al. 1988; Perkins et al. 1988). Footprinting analysis of these TATA-less promoters with eTFIID revealed that the factor did not appear to induce upstream DNase I protection or hypersensitivity in the −25 to −30 region (relative to the RNA start site) where TATA boxes are typically found (Fig. 3B). eTFIID did bind, however, to the promoters in a region that extends from the vicinity of the RNA start site to downstream as far as about +30 to +35. In contrast, dTBP did not appear to bind to these TATA-less promoters (data not shown). In addition, similar patterns of DNase I protection and hypersensitivity were observed with other
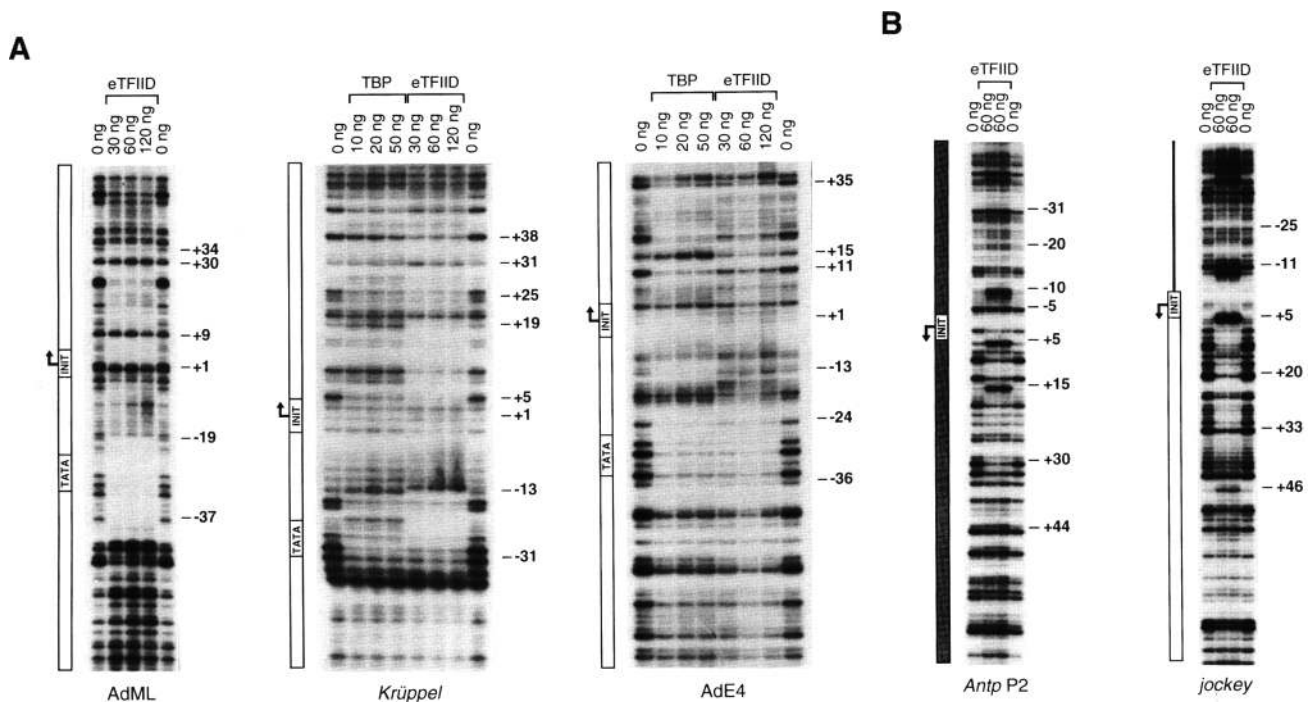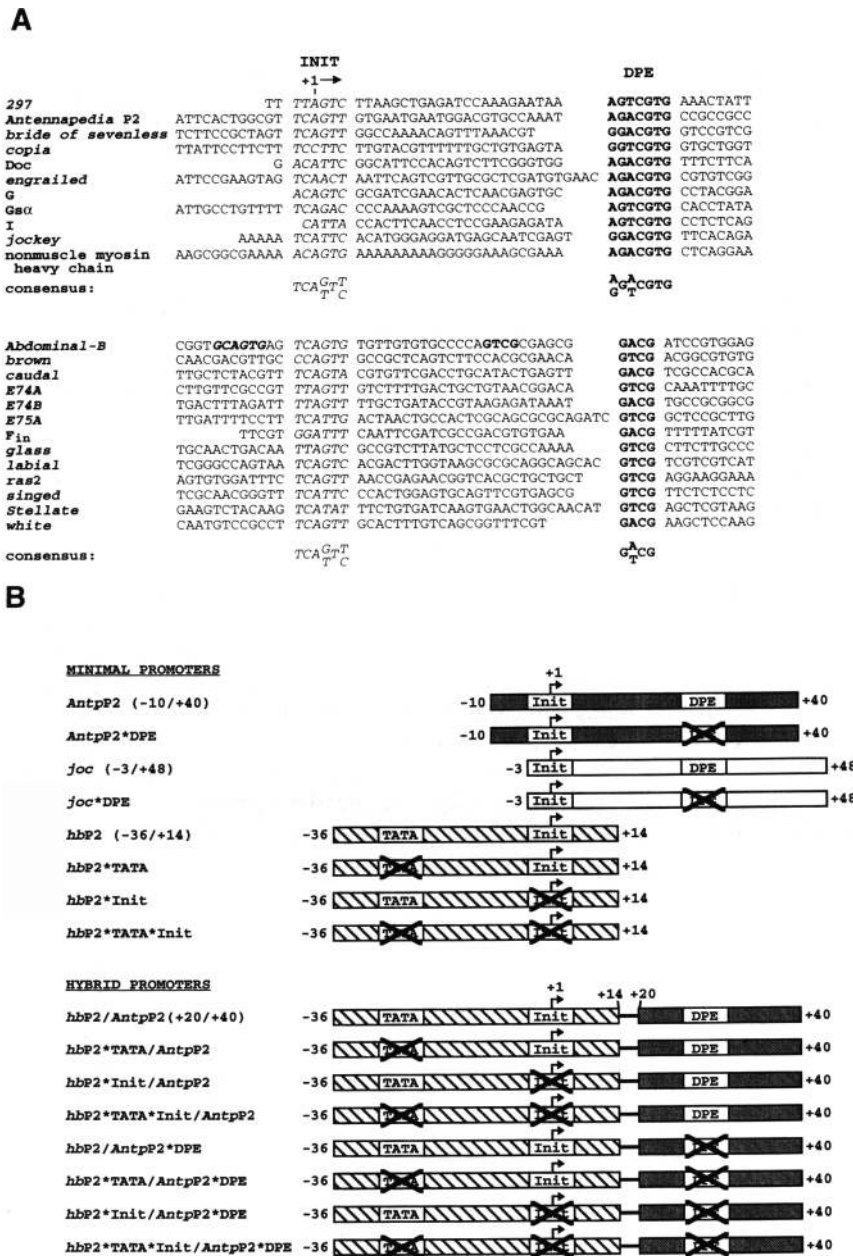

Figure 3. eTFIID binds to an extended region of both TATA-box-containing and TATA-less promoters. The positions of the TATA boxes, transcription initiation sites, protected regions, and hypersensitive sites are indicated, relative to the transcription start site as +1. The amount of TBP or eTFIID complex that was used in each footprinting reaction is indicated. (*A*) DNase I footprinting analysis of TBP and eTFIID with TATA-box-containing promoters: AdML, *Kr*, and AdE4. (*B*) DNase I footprinting analysis of eTFIID with TATA-less promoters: *Antp*P2 and *joc*.

A

B

Figure 4. The DPE of *Drosophila* TATA-less promoters. (*A*) Nucleotide sequence of the minimal promoter region of DPE-containing *Drosophila* promoters. Promoters that possess the DPE consensus are listed at the *top*; promoters that contain the $G^A/_TCG$ subset of the DPE consensus are listed at the *bottom*. This list of promoters is not meant to be comprehensive. Promoters from retrotransposons and LINE-like elements include *297* (Inouye et al. 1986), *copia* (Sinclair et al. 1986), Doc (Schneuwly et al. 1987), F (Di Nocera et al. 1983), G (Di Nocera 1988), I (Fawcett et al. 1986), and *joc* (Mizrokhi et al. 1988). Promoters from cellular genes include *Abdominal-B* (DeLorenzi et al. 1988), *AntpP2* (Perkins et al. 1988), *bride of sevenless* (Hart et al. 1993), *brown* (Dreesen et al. 1988), *caudal* (Mlodzik and Gehring 1987), *E74* (Thummel et al. 1989; Burtis et al. 1990), *E75* (Segraves and Hogness 1990), *engrailed* (Soeller et al. 1988), *glass* (Moses et al. 1989), Gsα (guanine nucleotide-binding protein; Quan and Forte 1990), *labial* (Mlodzik et al. 1988), nonmuscle myosin heavy chain (Ketchum et al., 1990), *ras2* (Bishop et al. 1988), *singed* (Paterson and O'Hare 1991), *Stellate* (Livak 1990), and *white* (O'Hare et al. 1984). The DPE sequences are given in boldface type, whereas the initiation site consensus sequences are indicated in italics. The *Abdominal-B* promoter appears to contain two initiation site consensus sequences (in italics) as well as two corresponding downstream consensus sequences (in boldface). Consensus sequences are given below each promoter element. (*B*) Schematic representations of minimal and hybrid promoters used in the functional analysis of the DPE. The specific DNA sequences in these promoters are described in Materials and methods. The TATA box element, Init, and DPE are shown in boxes. A cross over an element indicates that the native DNA sequence was disrupted by clustered point mutagenesis. The numbers denote positions relative to the major transcription start site (+1), which is indicated by an arrow. In the *hb*P2/*Antp*P2 hybrid promoters, the spacing between the *hb*P2 transcription initiation site and the *Antp*P2 DPE is identical to that in the wild-type *Antp*P2 promoter.

Drosophila TATA-less promoters, which include those from the *Ultrabithorax* and *E74* genes (data not shown). These data collectively suggest that the critical interactions in the binding of TFIID to TATA-less promoters are downstream of the normal position of the TATA box.

Identification and functional analysis of a downstream promoter element

Because the TATA box provides a high-affinity interaction of TFIID with TATA-box-containing promoters, we considered the possibility that there may be a functionally analogous downstream basal promoter element that

might promote an alternate mode of TFIID binding to TATA-less promoters. Consistent with this notion, previous studies of *Drosophila* TATA-less promoters have revealed a requirement for sequences ~20–40 nucleotides downstream of the RNA start site for efficient transcription. These TATA-less promoters include those from *Drosophila* retrotransposons (Mizrokhi and Mazo 1990; Jarrell and Meselson 1991; Arkhipova and Ilyin 1991; McLean et al. 1993) as well as promoters of genes encoding regulatory factors (Biggin and Tjian 1988; Perkins et al. 1988; Soeller et al. 1988; Thummel 1989). Hence, the available data were consistent with the hypothesis that there may be a downstream promoter ele-

ment in the +20 to +40 region that is important for binding of TFIID.

We thus inspected the downstream sequences of a few *Drosophila* TATA-less promoters and identified a conserved $G^A/_TCG$ motif at approximately +30. This analysis was subsequently expanded to include the 252 independent *D. melanogaster* promoter sequences (comprising 129 TATA-containing and 123 TATA-less promoters) that were compiled and studied by Arkhipova (1995). By using this *Drosophila* promoter data base (kindly provided by I. Arkhipova, Harvard University, Cambridge, MA), we identified 24 *Drosophila* promoters that contained the $G^A/_TCG$ sequence motif located ~25–30 nucleotides downstream of a *Drosophila* initiation site consensus sequence (Fig. 4A). None of these 24 promoters contained a TATA-box-like sequence in the −25 to −30 region upstream of the transcription start site. Moreover, 11 of the 24 sequences contained an extended downstream consensus sequence of $^A/_GG^A/_TCGTG$ (Fig. 4A, top sequences), and it is notable that two segments (CGTG and ACGY) within this extended consensus were identified previously by Arkhipova (1995) as motifs that were overrepresented in TATA-less promoters. Furthermore, re-examination of earlier studies of *Drosophila* TATA-less promoters revealed that a significant decrease in transcriptional activity accompanied the loss (typically by deletion) of the downstream consensus sequence (see, e.g., Perkins et al. 1988; Soeller et al. 1988; Thummel 1989; Jarrell and Meselson 1991; McLean et al. 1993). Hence, based on these data, we have tentatively designated the downstream consensus sequence, $^A/_GG^A/_TCGTG$, as the DPE.

To study the possible function of the DPE in TFIID

binding and transcriptional activity of TATA-less promoters, we constructed wild-type and DPE–mutant versions of the *Antp*P2 and *joc* promoters, as outlined at the top of Figure 4B. Both of these promoters contain an extended DPE consensus sequence (Fig. 4A). DNase I footprinting analysis of these defined, minimal promoters revealed that mutation of the DPE resulted in a significant decrease in the binding of eTFIID to the *Antp*P2 promoter as well as to the *joc* promoter (Fig. 5A). It is particularly notable that disruption of the DPE of either promoter resulted in a significant decrease in the interaction of TFIID not only in the downstream region, but also in the vicinity of the RNA start site. These results indicate that the presence of the DPE is important for the binding of TFIID to the TATA-less *Antp*P2 and *joc* promoters.

We then examined the transcriptional activity of the wild-type and DPE–mutant versions of the *Antp*P2 and *joc* promoters by in vitro transcription analysis. As shown in Figure 5B, disruption of the DPE causes a 20- to 100-fold reduction in the transcriptional activity of the promoters. Therefore, the combined footprint and transcriptional data indicate that disruption of the DPE causes a reduction in the binding of TFIID to the TATA-less promoters and a consequent loss of transcriptional activity.

The downstream promoter element can compensate for disruption of the TATA box but not alteration of the initiation site sequence

Next, we investigated whether the DPE in its normal downstream position can function to substitute for the

Figure 5. The DPE appears to be important for the binding of TFIID and the transcriptional activity of the *Drosophila* *Antp*P2 and *joc* promoters. The promoter constructions used are depicted in Fig. 4B. (A) DNase I footprint analysis of the binding of eTFIID to wild-type and DPE-mutated *Antp*P2 and *joc* promoters. The positions of the initiation site sequences (INIT) and the DPEs are indicated. A cross over the DPE element indicates that the native DNA sequence was disrupted by clustered point mutagenesis. The amount of TBP or eTFIID complex that was used in each footprinting reaction is indicated. (B) In vitro transcription analysis of wild-type and DPE-mutated *Antp*P2 and *joc* promoters. Standard transcription reactions were performed with the indicated template DNAs (500 ng) and *Drosophila* SK nuclear extract (100 μg total protein), and the resulting transcripts were subjected to primer extension analysis. The reverse transcription products are indicated by brackets. Transcriptional activity is reported as relative to that of the wild-type version of each promoter.
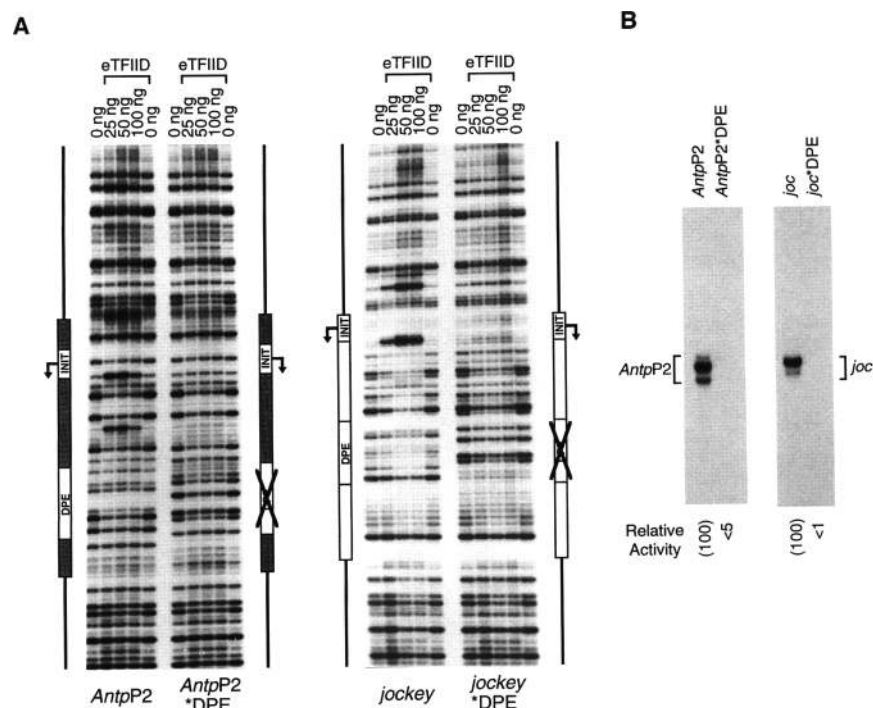
upstream TATA box motif of a TATA-containing promoter. In these experiments, we characterized wild-type and mutant versions of the minimal *Drosophila hunchback* P2 (*hb*P2) promoter along with a series of *hb*P2/*Antp*P2 hybrid promoters, which are outlined in Figure 4B. The *hb*P2 promoter contains a TATA box (TATAAA) upstream of dual, overlapping consensus initiation site sequences at +1 and +5, from which transcription initiates both in vitro and in vivo. The minimal *hb*P2 promoter contains the TATA box and initiation site sequence of the *hb*P2 promoter while the hybrid *hb*P2/*Antp*P2 construction contains the TATA box and initiation site sequence of the *hb*P2 promoter fused to the DPE of the *Antp*P2 promoter, with the identical spacing between the DPE and the *hb*P2 transcription start site (+1) as in the wild-type *Antp*P2 promoter. Mutant variants of these promoters with clustered point mutations in the TATA box (*TATA), initiation site sequences (*Init), and the DPE (*DPE) were also examined.

This set of promoter constructions was then subjected to in vitro transcription analysis (Fig. 6A). With the *hb*P2 minimal promoter, disruption of the TATA box element resulted in a reduction of the transcriptional activity to ~13% of that of the corresponding wild-type promoter (Fig. 6A, cf. lanes 1 and 2). In contrast, with the *hb*P2/*Antp*P2 hybrid promoter, the *Antp*P2 DPE partially compensated for the disruption of the *hb*P2 TATA box, as the *TATA version exhibited ~53% of the activity of the corresponding TATA-containing promoter (Fig. 6A, cf. lanes 5 and 6). By comparison, a mutant version of the *Antp*P2 DPE was unable to restore activity of the disrupted TATA box (Fig. 6A, lanes 9,10). These results suggest that the DPE can compensate for the loss of the TATA box. On the other hand, the DPE did not restore transcriptional activity to promoters in which the initiation site sequence was altered (Fig. 6A, lanes 3,7,11). It thus appears that the DPE can function in the absence of the *hb*P2 TATA box but not in the absence of the *hb*P2 initiation site sequence.

It is also useful to compare the usage of the +1 versus +5 RNA start sites of the *hb*P2 promoter (or the *hb*P2/*Antp*P2*DPE promoter) in reactions performed with the *hb*P2*TATA/*Antp*P2 promoter (Fig. 6A, lane 6). In the absence of the TATA box and in the presence of a compensatory DPE, the +1 site was preferentially transcribed relative to the +5 site, whereas with the wild-type *hb*P2 promoter (or the *hb*P2/*Antp*P2*DPE promoter), the +5 site was slightly favored relative to the +1 site (Fig. 6A, lanes 1,9). These findings therefore suggest that the preference for the +1 site with the *hb*P2*TATA/*Antp*P2 promoter is attributable to a preferred positioning of the DPE relative to the +1 RNA start site, because in that construction, the nucleotide spacing between *Antp*P2 DPE and the +1 start site, but not the +5 site, is identical to that in the wild-type *Antp*P2 promoter.

We then carried out a DNase I footprinting analysis of the *hb*P2/*Antp*P2 constructions to see whether the binding of TFIID correlated with the transcriptional activity of the promoters (Fig. 6B). With the *hb*P2/*Antp*P2 and

the *hb*P2/*Antp*P2*DPE promoters, both of which contain intact TATA box elements, the DNase I footprints were nearly identical and extended from about −36 to +40 relative to the RNA start site (Fig. 6B, left). Thus, in the presence of a TATA box, the DPE had little effect on the binding of TFIID throughout the promoter region. In contrast, with the TATA-less promoters, the DPE had a significant effect. As shown in Figure 6B (right), the binding of TFIID to the *hb*P2*TATA/*Antp*P2 promoter was greater than that to the *hb*P2*TATA/*Antp*P2*DPE promoter. Hence, the TFIID footprinting results correlate well with the transcription data. It is also notable that there was no apparent DNase I footprint in the region of the mutated TATA box of the *hb*P2*TATA/*Antp*P2 promoter. Instead, the footprint of this TATA⁻ Init⁺ DPE⁺ promoter resembled the footprints of the *Antp*P2 and *joc* TATA-less promoters (Fig. 5A).

Because a significant decrease in transcription was observed upon mutation of the Init sequences (Fig. 6A), we additionally tested the binding of TFIID to promoters containing the DPE along with wild-type versus mutant initiation site (Init) regions (Fig. 6C). DNase I footprinting of the *hb*P2/*Antp*P2 versus the *hb*P2*Init/*Antp*P2 promoters (Fig. 6C, left) and the *hb*P2*TATA/*Antp*P2 versus the *hb*P2*TATA*Init/*Antp*P2 promoters (Fig. 6C, right) revealed that the Init sequences contribute to the binding of TFIID to the downstream promoter region (comprising the Init and DPE sequences) of the promoters. Hence, the transcription and footprinting data collectively suggest that interactions of TFIID with the Init and DPE sequences can be important for basal transcription by RNA polymerase II.

Systematic mutational analysis of the DPE

The DPE was initially defined on the basis of the conservation of the $^A/_GG^A/_TCGTG$ sequence motif (Fig. 4A), and we therefore sought to determine, in a systematic manner, the importance of the DPE consensus sequence in basal RNA polymerase II transcription. To this end, we constructed and analyzed a series of clustered triple point mutations in the *Drosophila joc* promoter from +20 to +40 relative to the RNA start site. Transcriptional analysis of these mutant promoters revealed that alteration of the core of the DPE consensus, but not the flanking sequences, resulted in a sharp decrease in the efficiency of transcription from the *joc* promoter (Fig. 7). Hence, these experiments reveal the importance of the DPE consensus sequence for basal transcription and further support the hypothesis that the DPE is a distinct, basal downstream promoter element.

Discussion

In this work we have identified and characterized a conserved basal DPE that is present in a subset of *Drosophila* TATA-less promoters. The DPE, in conjunction with the initiation site sequence, provides a binding site for TFIID in the absence of a TATA box. In the context of our current knowledge of TFIID and basal transcrip-
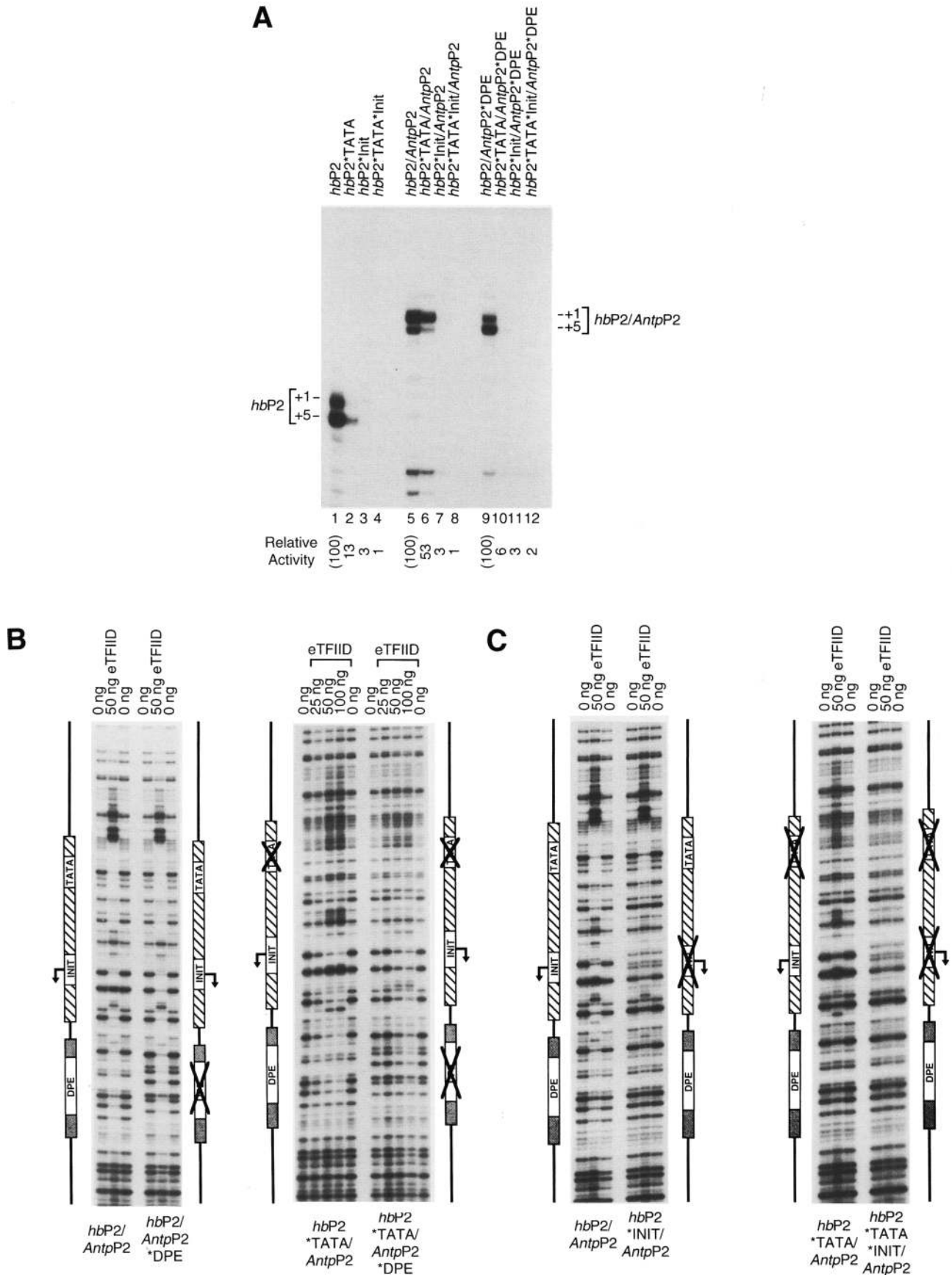
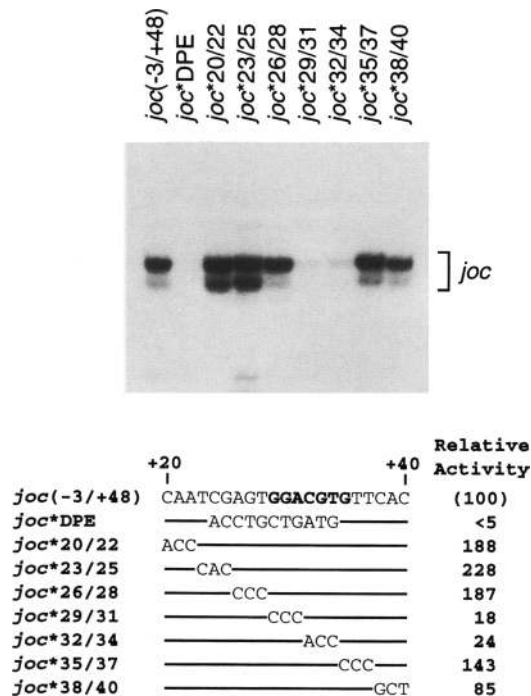**Figure 6.** (*See facing page for legend.*)

|	+20	+40	Relative Activity
joc(-3/+48)	CAATCGAGT**GGACGTG**TTCAC		(100)
joc*DPE	——————ACCTGCTGATG————		<5
joc*20/22	ACC—————————————		188
joc*23/25	————CAC—————————		228
joc*26/28	—————————CCC———————		187
joc*29/31	—————————————CCC————		18
joc*32/34	———————————————ACC————		24
joc*35/37	—————————————————CCC——		143
joc*38/40	———————————————————GCT		85

Figure 7. Systematic clustered point mutational analysis of the DPE in the *joc* promoter. Standard transcription reactions were performed with the indicated template DNAs (500 ng) and *Drosophila* SK nuclear extract (100 µg total protein), and the resulting transcripts were detected by primer extension analysis. The reverse transcription reaction products are indicated by brackets. The *joc* promoter sequence from +20 to +40 relative to the major transcription start site is shown with corresponding nucleotide substitutions below. Lines indicate unchanged sequences. The DPE consensus sequence is shown in bold. Transcriptional activity is reported as relative to that of the wild-type *joc*(-3/+48) promoter.

tion, these studies support and clarify the model that there are different modes by which TFIID interacts with promoters. Figure 8 depicts two possible models for the binding of TFIID to basal promoter sequences. The top panel of Figure 8 shows a promoter with a strong TATA box element. In such TATA-driven promoters [e.g., the AdML, AdE4, *Drosophila Kr*, and *Drosophila hb*P2 promoters (Figs. 3A and 6B)], TFIID binds to the TATA motif and often (as with the AdML, *Kr*, and *hb*P2 promot-
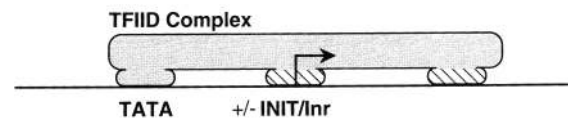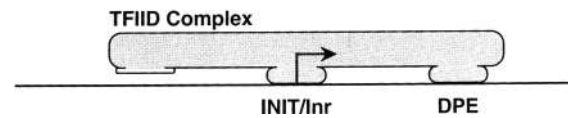

Figure 8. A simple model for interactions of the TFIID complex with TATA-containing and TATA-less promoters. The hatched regions indicate that the corresponding TFIID–DNA contacts are observed with some promoters.

ers), but not always (as with the AdE4 promoter), interacts closely with downstream regions in a manner that leads to distinct protection from DNase I digestion in footprinting assays. The bottom panel of Figure 8 depicts a TATA-less promoter that contains a DPE. With these promoters [e.g., the *Drosophila Antp*P2 and the *Drosophila joc* promoters (Figs. 3B and 5)], a TFIID-mediated footprint is not observed in the typical location of the TATA box (about -25 to -30). In contrast, TFIID appears to interact closely with a downstream region that encompasses the initiation site sequence and the DPE. It should be also noted, parenthetically, that the two models shown in Figure 8 are meant to represent only a subset of the potential interactions of TFIID with basal promoters.

The DPE

In this work we characterized two TATA-less promoters, *Antp*P2 and *joc*, that possess the DPE consensus sequence, $^A/_GG^A/_TCGTG$, which is shown in Figure 4A (top). This DPE consensus is likely to be an optimized, high affinity recognition site for the downstream binding of TFIID to TATA-less promoters. As with TATA boxes and Inr elements, it seems probable that other promoters will contain related versions of the DPE sequence to which TFIID would interact with varying affinity. Thus,

Figure 6. The *Antp*P2 DPE can partially compensate for the loss of the TATA box of the *hb*P2 promoter. The promoter constructions used in this figure are shown in Fig. 4B. (*A*) Transcriptional analysis of wild-type and mutant *hb*P2 and *hb*P2/*Antp*P2 hybrid promoters. Standard transcription reactions were performed with the indicated template DNAs (500 ng) and *Drosophila* SK nuclear extract (100 µg total protein), and the resulting transcripts were detected by primer extension analysis. The reverse transcription reaction products are indicated by brackets. Transcriptional activity is reported as relative to that of the wild-type version of each promoter. (*B*) DNase I footprint analysis of the binding of eTFIID to wild-type and mutant *hb*P2/*Antp*P2 hybrid promoters. The positions of the TATA box elements, initiation site sequences (INIT), and the DPEs are indicated. A cross over the TATA box or DPE element indicates that the native DNA sequence was disrupted by clustered point mutagenesis. The amount of eTFIID complex that was used in each footprinting reaction is indicated. (*C*) DNase I footprint analysis of the binding of eTFIID to wild-type and mutant *hb*P2/*Antp*P2 hybrid promoters. The positions of the TATA box elements, INIT, and DPEs are indicated. A cross over the TATA box or INIT element indicates that the native DNA sequence was disrupted by clustered point mutagenesis. The amount of eTFIID complex that was used in each footprinting reaction is indicated.

although we had identified only 11 DPE-containing promoters out of a data base of 252 promoters, there are probably many other promoters, perhaps some of those shown in the lower portion of Figure 4A, that contain functionally important sequences that resemble the DPE consensus.

Why has the DPE been identified only in TATA-less promoters? It is possible, for instance, that the presence of a TATA box eliminates the requirement for a DPE. This postulate is supported by our results with the *hb*P2/*Antp*P2 hybrid promoters (Fig. 6). When the intact *hb*P2 TATA box was present, the addition of the DPE had almost no detectable effect upon either the transcriptional activity of the promoter (Fig. 6A, cf. lanes 1, 5, and 9) or the binding of TFIID throughout the promoter region both upstream and downstream of the RNA start site (Fig. 6B, left). In this context, it should also be noted that some of the DPE-containing promoters listed in Figure 4A were those of LINE-like elements, which are required to have internal promoters and thus could not have upstream TATA box motifs. Hence, the DPE might be particularly useful in promoters that cannot accommodate TATA box elements. Yet, on the other hand, the DPE appears to be compatible with the TATA box; thus, there may be promoters that contain both the DPE and the TATA box element. One example of such a promoter may be the TATA box-containing *Drosophila hsp70* promoter, which has been shown to contain a downstream GTCG motif that is protected from hydroxyl radical footprinting by TFIID (Purnell et al. 1994).

It is also important to consider whether DPE-like elements are present in the promoters of organisms other than *Drosophila*. Given that the basal transcription process is conserved from yeast to humans, it seems likely that there are DPEs that are functionally analogous to the DPE in other eukaryotes. Yet, by comparison with initiation site sequences, which exhibit a much more distinct consensus in *Drosophila* (Hultmark et al. 1986; Purnell et al. 1994; Arkhipova 1995) than in other eukaryotes, it is similarly possible that the DPE sequences in *Drosophila* promoters may adhere more closely to the DPE consensus than related DPE-like motifs in other species. Hence, the identification of DPE motifs in another organism, such as humans, may be more difficult than in *Drosophila*.

We also found that purified eTFIID, but not TBP, bound to the *Antp*P2 and *joc* TATA-less promoters, as determined by DNase I footprinting analysis. We therefore considered the possible relation between our findings and earlier observations (Sypes and Gilmour 1994; Verrijzer et al. 1994, 1995) that the 150-kD TAF subunit of dTFIID (TAF$_{II}$150) appears to be involved in the binding of TFIID downstream of the RNA start site of a *Drosophila hsp70* promoter and the AdML promoter. Inspection of the sequences of the *hsp70* and AdML promoters did not reveal any close match to the DPE consensus. The AdML promoter does not appear to have any downstream sequences with any resemblance to the DPE, whereas the *hsp70* promoter contains a downstream GTCG sequence that corresponds to a portion of

the DPE. Thus, it is difficult to attribute the binding of purified, recombinant TAF$_{II}$150 to the AdML downstream sequences (Verrijzer et al. 1994) to the presence of a DPE-like sequence. In addition, the presence of the consensus TATA box motifs in the AdML and *hsp70* promoters may have also contributed to the downstream interactions of either TFIID or TBP–TAF$_{II}$250–TAF$_{II}$150 complexes with those promoters (Sypes and Gilmour 1994; Verrijzer et al. 1995). In the future, it will be important to investigate in greater detail the possible interactions of the TAFs with the DPE in the context of a TATA-less promoter.

Perspectives

There are multiple mechanisms of basal transcription from TATA-containing and TATA-less promoters. The combined use of the consensus initiation site sequence along with the DPE, as suggested by the results of this study, is likely to be only one of several modes of transcription from TATA-less promoters. In this context, the identification of the DPE should be a useful contribution to our understanding of basal transcription. These results also support the hypothesis that there are three potential recognition sites for TFIID in core promoters: the TATA box, the initiation site sequence/Inr, and the DPE. Further studies of purified transcription factors and core promoter elements should eventually illuminate the varied mechanisms of basal transcription by RNA polymerase II.

Materials and methods

Nomenclature

Transcription factors IIA through IIS are abbreviated as TFIIA, TFIIB, and so on. The TATA box-binding subunit of the TFIID complex is referred to as TBP. The epitope-tagged versions of *Drosophila* TBP and TFIID are referred to as eTBP and eTFIID. The purified basal transcription factors used in this study were synthesized in *Escherichia coli*. A lowercase d or h designates the original cDNA source of the recombinant factors as either *Drosophila melanogaster* or human. The 34- and 56-kD subunits of human TFIIE are referred to as hTFIIE34 and hTFIIE56, respectively. The 30- and 74-kD subunits of human TFIIF are referred to as hTFIIF30 and hTFIIF74, respectively. The complex of the 14- and 48-kDa subunits of *Drosophila* TFIIA is referred to as dTFIIA. Alternate nomenclature for the basal transcription factors (for review, see Zawel and Reinberg 1992) is as follows: TFIIA (STF), TFIIB (α, FA, factor e), TFIID (τ, factor d, BTF-1), TFIIE (ϵ, factor a), TFIIF (RAP30/74, $\beta\gamma$, factor 5, FC, BTF-4, factor g), TFIIH (δ, factor b, BTF-2), and TFIIS (S-II).

To maintain the integrity of the original definition of the "initiator" (Inr) as a functional basal promoter element that can direct transcription in the absence of an upstream TATA box (Smale and Baltimore 1989; Smale 1994), we refer to the DNA sequence that encompasses the RNA start site as the initiation site sequence (Init) rather than as the initiator (Inr) element. Unlike Inr elements, we have found that the *Drosophila* initiation site sequences alone are insufficient to direct a detectable level of basal transcription (see, e.g., Kerrigan et al. 1991).

Plasmids and synthetic oligonucleotides

The P-element injection vector pCaSpeR–hs–eTBP was constructed as follows. This vector contains a *Drosophila hsp70* promoter that directs the transcription of a modified *Drosophila* TBP cDNA that encodes dTBP with a FLAG epitope (IBI/Kodak) followed by 6 histidine residues at the amino-terminus. Two oligonucleotides were synthesized for PCR-directed mutagenesis of the dTBP cDNA to generate amino-terminal tags, as well as restrictions sites for cloning. The amino-terminal primer (5′eTBP; 5′-CTGGATCCGTTAACGCCATGGACTACAAGG-ATGACGATGACAAGCACCATCACCACCATCATGACCA-AATGCTAAGCCCCAACTTC-3′) contained *Bam*HI, *Hpa*I, and *Nco*I sites that were followed by a consensus ribosomal initiation sequence (Kozak 1986) upstream of the sequences encoding the FLAG epitope, 6 histidines, and amino acids 2–9 of dTBP. The carboxy-terminal primer (3′eTBP; 5′-GAGAATTC-TCTAGACGTTATGACTGCTTCTTGAACTTCTT-3′) contained *Eco*RI and *Xba*I sites followed by sequences complementary to amino acids 353–347 of dTBP as well as the ochre stop codon. A dTBP cDNA [template pdTBP353 (Hoey et al. 1990)] was PCR-amplified with the above primers, digested with *Bam*HI and *Eco*RI, and then subcloned into the corresponding restriction sites of pBlueScript SK (Stratagene) to give pBS–eTBP. To ensure that spurious mutagenesis had not occurred during PCR, the entire TBP cDNA and flanking regions of pBS–eTBP were sequenced. The *Hpa*I–*Xba*I fragment of pBS–eTBP containing the dTBP cDNA was then subcloned into *Hpa*I and *Xba*I sites of the P-element transformation vector pCaSpeR–hs [containing a 452-bp *Xba*I-*Xmn*I fragment of the *hsp70* promoter and a 525-bp *Eco*RI–*Pst*I fragment of the *hsp70* 3′ region, inserted into pCaSpeR (Pirrotta 1988); provided by T. Ip (University of Massachusetts Medical Center, Worcester)], to give pCaSpeR–hs–eTBP.

A series of plasmids containing minimal promoters were constructed by insertion of double-stranded oligonucleotides into the *Xba*I and *Pst*I sites of the polylinker of pUC119 (see Fig. 4B). *hb*P2 wild-type and TATA (*TATA) and/or initiation site (*Init) mutant promoter sequences were described previously (George et al. 1995). In the *TATA mutant *hb*P2 promoters, the native TATA box sequence TATATAAA was replaced with ACGTCCGT. In the *Init mutant *hb*P2 promoters, the native Init sequence, CAGTCA, was replaced with GTGGTT. The following sequences were inserted into the *Xba*I and *Pst*I sites of pUC119 to create *joc* and *Antp*P2 minimal promoters (these plasmids were generously provided by C. George, University of California, San Diego, La Jolla): 5′-ATCATTCGCATGG-GAGATGAGCAAT**CGAGTGGACGT**GTTCAC-AGAAGTC-GC-3′ (*joc* − 3/ + 48) and 5′-CACTGGCGTTCAGTTGTGAAT-GAATGGACGTGC**CAAATAGACGT**GCCGCC-3′ (*Antp*P2 − 10/ + 40). The boldface nucleotides of *Antp*P2 and *joc* were replaced by the sequence 5′-ACCTGCTGATG-3′ to generate *Antp*P2*DPE and *joc**DPE minimal promoters. The clustered triple point mutant promoters used in Figure 7 were constructed in an analogous fashion. The *hb*P2/*Antp*P2 hybrid promoter series was constructed by inserting double-stranded oligonucleotides identical to nucleotides + 21 to + 40 relative to the start site of either *Antp*P2 or *Antp*P2*DPE into the *Pst*I site immediately downstream of each of the *hb*P2 minimal promoters. In these hybrid promoters, the spacing between the *hb*P2 transcription initiation site and the *Antp*P2 DPE is identical to that in the wild-type *Antp*P2 promoter.

Purification of transcription factors

Basal transcription factors dTBP, dTFIIB, hTFIIE34, hTFIIE56, hTFIIF30, and hTFIIF74 were synthesized in *E. coli* and purified

to >90% homogeneity as described previously (Tyree et al. 1993). dTFIIA [plasmids pAR3038–dTFIIA–S (Yokomori et al. 1994) and pAR3038–dTFIIA–L (Yokomori et al. 1993) were kindly provided by K. Yokomori and R. Tjian (both at University of California, Berkeley)] was prepared as follows. First, the 14- and the 48-kD subunits of dTFIIA were each synthesized separately in *E. coli*, solubilized in urea, and then combined and renatured together as described previously (Yokomori et al. 1994). The resulting dTFIIA was further purified by successive chromatography on Q Sepharose and Superdex 200 resins (Pharmacia). TFIIH was partially purified from *Drosophila* embryos to the extent that the preparation was deficient in RNA polymerase II and the other basal transcription factors (L. Lira-DeVito and J.T. Kadonaga, unpubl.). RNA polymerase II was purified to >90% homogeneity from calf thymus (George et al. 1995).

P-element-mediated germ-line transformation

Germ-line transformants were obtained by injecting pCaSpeR–hs–eTBP DNA (1 mg/ml) with the P-transposase helper plasmid δ2,3 (100 μg/ml) into *white*⁻ embryos (*yw^{67c23}/yw^{67c23}*) (Rubin and Spradling 1982). Of the 505 embryos injected, 199 completed embryogenesis and 101 developed to the adult stage. Nine independent transformed lines were obtained from the surviving G₀ adults. Heterozygotes were crossed with balancer strains (CyO or TM3) to generate homozygous transformed flies, and one hs–eTBP strain was expanded into population cages for large-scale embryo collections. Although the *hsp70* promoter was used for the transcription of the eTBP cDNA, the flies were grown and maintained at 25°C under standard conditions that are employed for large-scale growth of *Drosophila*. Hence, the basal level of transcription from the uninduced *hsp70* promoter provided the appropriate level of expression of the eTBP cDNA for the subsequent purification of eTFIID. Under these conditions, we have found that the transformed flies stably express eTBP.

Purification of eTFIID from transgenic *Drosophila* embryos

All operations were performed at 4°C. Nuclear extracts were prepared with hs–eTBP embryos collected between 0 and 12 hr after fertilization and stored for ≤3 days at 4°C. Embryos were harvested (typically 50–60 grams after dechorionation), and nuclear extracts were prepared according to Wampler et al. (1990), except that the 2.26 M ammonium sulfate pellet was resuspended in HEGN buffer containing 0.1 M KCl [HEGN is 25 mM HEPES, K⁺ (pH 7.6), 0.1 mM EDTA, 10% (vol/vol) glycerol, 0.1% (vol/vol) NP-40, 1 mM dithiothreitol, 1 mM benzamidine–HCl, 1 mM sodium metabisulfite, 0.1 mM phenylmethylsulfonyl fluoride] and then desalted by dialysis against HEGN buffer containing 0.1 M KCl. This mixture was subjected to centrifugation in a Sorvall SS-34 rotor at 10,000 rpm for 10 min to remove insoluble material. The extract was fractionated on a phosphocellulose P-11 (Whatman) column [8-ml column volume, 0.5 ml/min flow rate, 1.5 ml fraction size, full-scale absorbance at 280 nM = 2]. The extract was applied to the column, which had been extensively equilibrated in HEGN buffer containing 0.1 M KCl, and washed with three column volumes of the same buffer. The wash was followed with 0.3 M 0.45 M, and 0.85 M KCl elution steps (three column volumes each), and the fractions containing the majority of the protein were pooled for each step. The 0.45–0.85 M KCl fraction was diluted with an equal volume of HEGN buffer without KCl. Each P-11 fraction was incubated with 50 μl (bed volume) of anti-FLAG M2 agarose resin (IBI/Kodak) for 8–12 hr by rotation. After three washes in HEGN buffer containing 0.1 M KCl [(10 ml each) fractions were

centrifuged for 3 min at 2000 rpm in SS-34, and all but 500 µl of buffer was carefully removed and resuspended], the agarose resin was transferred to a microcentrifuge spin column (InVitrogen) and residual buffer was removed by centrifugation at 2000 rpm for 1–2 min. The bound protein complexes were eluted from the anti-FLAG M2 resin by incubation with 100 µl of HEGN buffer containing 12.5 mM MgCl$_2$, 0.1 M KCl, and 0.2 mg/ml of the FLAG peptide (DYKDDDDK; IBI/Kodak) for 1 h. The eluate was collected by centrifugation at 2000 rpm for 1–2 min. The majority of the eTFIID was found in the 0.45–0.85 M KCl P-11 fraction, and this material was used for the experiments. eTBP Western units in Table 1 were generated by Western blot analysis. Protein fractions were resolved by 9% SDS-PAGE, transferred to nitrocellulose, and probed with anti-dTBP polyclonal antisera. Bands were detected with ^{125}I-labeled protein A (ICN Pharmaceuticals), and eTBP was quantitated by using a PhosphorImager (Molecular Dynamics or Fuji).

In vitro transcription analysis

Transcription reactions were carried out as described previously (Wampler et al. 1990; Tyree et al. 1993) by using supercoiled DNA templates (100 or 500 ng, as indicated) with either purified basal transcription factors (Tyree et al. 1993) or an SK (Soeller–Kornberg) extract from *Drosophila* embryos (Soeller et al. 1988) in a total volume of 25 µl. Ribonucleoside 5'-triphosphates were either added following a 30 min preincubation at 21°C (purified factors) or directly to the reaction media in the absense of a preincubation (extracts). The resulting transcripts were subjected to primer extension analysis as described previously (Kadonaga 1990; Wampler et al. 1990). Transcripts synthesized from the minimal promoters were analyzed with the M13 reverse sequencing primer (5'-AGCGGATAACAATTTCACACAGGA-3'), which is complementary to sequences in the pUC plasmid. Transcription and primer extension analysis with AdE4 and AdML promoters were performed as described previously (Kerrigan et al. 1991; Tyree et al. 1993). Quantitation of the in vitro-synthesized RNA was carried out with a PhosphorImager (Molecular Dynamics or Fuji). All experiments were performed a minimum of two times (but typically at least four times) to ensure reproducibility of the data.

DNase I footprint analysis

DNase I footprint probes were prepared by PCR with primers flanking the promoter region. Either the upstream or downstream primer was 5'-^{32}P-labeled, as indicated in the figures. The PCR amplification products were gel-purified on 5% nondenaturing polyacrylamide gels to remove free primer and spurious reaction byproducts. Binding reactions for DNase I footprinting experiments were performed for 30 min at 25°C in a total volume of 50 µl and contained 12.5 mM HEPES, K$^+$ (pH 7.6), 50 mM KCl, 0.05 mM EDTA, 5% (vol/vol) glycerol, 0.05% (vol/vol) NP-40, 0.5 mM dithiothreitol, 2% polyvinyl alcohol, and ^{32}P-labeled probe (6000 cpm). Nonspecific competitor DNA was not included in the reactions. The labeled probes were then partially digested with DNase I, and the reaction products were purified and analyzed by electrophoresis on a 6% denaturing polyacrylamide gel. PCR primers used for amplification of minimal promoter DNA templates were the M13 universal (upstream; 5'-CGCCAGGGTTTTCCCAGTCACGA-3') and M13 reverse (downstream) sequencing primers. PCR primers used for full-length promoter footprinting probes were the following: 5'AdML (upstream; 5'-GTAGCGGTCGTTGTCCACTAGG-3') and 3'AdML (downstream; 5'-CCACCCTCAAAGGCATCACCG-3') for the AdML promoter template pLAX (Dynan and Tjian 1993); 5'E4 (upstream; 5'-GAATACAAGCTTGCAT-

GCCTG-3') and 3'E4 (downstream; 5'-CTAACAGTCAGCCTTACCAG-3') for the AdE4 promoter template pG$_5$E4T (Lin et al. 1988); 5'Kr (upstream; 5'-GCTTATTTGCGTGCGTGTGAG-3') and 3'Kr (downstream; 5'-CGCGGTTGTGTGTGTGCACAAC-3') for the *Drosophila Kr* promoter template pKr (Kerrigan et al. 1991); M13 reverse sequencing primer and 3'Joc (downstream; 5'-CGTCACCGTTATGCAGTGACTA-3') for the *Drosophila joc* promoter template pJC (Mizrokhi et al. 1988), and 5'Antp (upstream; 5'-CAGCCATTGGCAACATGCGATG-3') and 3'Antp (downstream; 5'-CGTCCGGTATTTAAGAACTGGG-3') for the *Drosophila AntpP2* promoter template pAntpP2–CAT (kindly provided by J. Posakony, University of California, San Diego, La Jolla).

Acknowledgments

References

Arkhipova, I.R. 1995. Promoter elements in *Drosophila melanogaster* revealed by sequence analysis. *Genetics* **139**: 1359–1369.

Arkhipova, I.R. and Y.V. Ilyin. 1991. Properties of promoter regions of mdg1 *Drosophila* retrotransposon indicate that it belongs to a specific class of promoters. *EMBO J.* **10**: 1169–1177.

Biggin, M.D. and R. Tjian. 1988. Transcription factors that activate the *Ultrabithorax* promoter in developmentally staged extracts. *Cell* **53**: 699–711.

Bishop, J.G. and V.G. Corces. 1988. Expression of an activated *ras* gene causes developmental abnormalities in transgenic *Drosophila melanogaster*. *Genes & Dev.* **2**: 567–577.

Buratowski, S. 1994. The basics of basal transcription by RNA polymerase II. *Cell* **77**: 1–3.

Buratowski, S., S. Hahn, P.A. Sharp, and L. Guarente. 1988. Function of a yeast TATA element-binding protein in a mammalian transcription system. *Nature* **334**: 37–42.

Buratowski, S., S. Hahn, L. Guarente, and P.A. Sharp. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**: 549–561.

Burtis, K.C., C.S. Thummel, C.W. Jones, F.D. Karim, and D.S. Hogness. 1990. The *Drosophila* 74EF early puff contains *E74*, a complex ecdysone-inducible gene that encodes two *ets*-related proteins. *Cell* **61**: 85–99.

Cavallini, B., I. Faus, H. Matthes, J.M. Chipoulet, B. Winsor, J.M. Egly, and P. Chambon. 1989. Cloning of the gene encoding the yeast protein BTF1Y, which can substitute for the human TATA box-binding factor. *Proc. Natl. Acad. Sci.* **86**: 9803–9807.

Chasman, D.I., J. Leatherwood, M. Carey, M. Ptashne, and R.D. Kornberg. 1989. Activation of yeast polymerase II transcription by herpes VP16 and GAL4 derivatives in vitro. *Mol. Cell. Biol.* **9:** 4746–4749.

Chen, J.-L., L.D. Attardi, C.P. Verrijzer, K. Yokomori, and R. Tjian. 1994. Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell* **79:** 93–105.

Chiang, C.-M., H. Ge, Z. Wang, A. Hoffmann, and R.G. Roeder. 1993. Unique TATA-binding protein-containing complexes and cofactors involved in transcription by RNA polymerases II and III. *EMBO J.* **12:** 2749–2762.

Conaway, R.C. and J.W. Conaway. 1993. General initiation factors for RNA polymerase II. *Annu. Rev. Biochem.* **62:** 161–190.

DeLorenzi, M., N. Ali, G. Saari, C. Henry, M. Wilcox, and M. Bienz. 1988. Evidence that the *Abdominal-B* r element function is conferred by a trans-regulatory homeoprotein. *EMBO J.* **7:** 3223–3231.

Di Nocera, P.P. 1988. Close relationship between non-viral retroposons in *Drosophila melanogaster. Nucleic Acids Res.* **16:** 4041–4052.

Di Nocera, P.P., M.E. Digan, and I.B. Dawid. 1983. A family of oligo-adenylate-terminated transposable sequences in *Drosophila melanogaster. J. Mol. Biol.* **168:** 715–727.

Dreesen, T.D., D.H. Johnson, and S. Henikoff. 1988. The brown protein of *Drosophila melanogaster* is similar to the white protein and to components of the active transport complexes. *Mol. Cell. Biol.* **8:** 5206–5216.

Dynan, W.S. and R. Tjian. 1983. Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. *Cell* **32:** 669–680.

Dynlacht, B.D., Hoey, T., and R. Tjian. 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66:** 563–576.

Fawcett, D.H., C.K. Lister, E. Kellet, and D.J. Finnegan. 1986. Transposable elements controlling I-R hybrid dysgenesis in *D. melanogaster* are similar to mammalian LINEs. *Cell* **47:** 1007–1015.

George, C.P., L.M. Lira-DeVito, S.L. Wampler, and J.T. Kadonaga. 1995. A spectrum of mechanisms for the assembly of the RNA polymerase II transcription preinitiation complex. *Mol. Cell. Biol.* **15:** 1049–1059.

Hart, A.C., S.D. Harrison, D.L. Van Vactor, Jr., G.M. Rubin, and S.L. Zipursky. 1993. The interaction of bride of sevenless is conserved between *Drosophila virilis* and *Drosophila melanogaster. Proc. Natl. Acad. Sci.* **90:** 5047–5051.

Hoey, T., B.D. Dynlacht, M.G. Peterson, B.F. Pugh, and R. Tjian. 1990. Isolation and characterization of the *Drosophila* gene encoding the TATA box binding protein, TFIID. *Cell* **61:** 1179–1186.

Hoey, T., R.O.J. Weinzierl, G. Gill, J.-L. Chen, B.D. Dynlacht, and R. Tjian. 1993. Molecular cloning and functional analysis of *Drosophila* TAF110 reveal properties expected of coactivators. *Cell* **72:** 247–260.

Hopp, T.P., K.S. Prickett, V. Price, R.T. Libby, C.J. March, P. Cerretti, D.L. Urdal, and P.J. Conlon. 1988. A short polypeptide marker sequence useful for recombinant protein identification and purification. *Biotechnology* **6:** 1205–1210.

Hultmark, D., R. Klemenz, and W. Gehring. 1986. Translational and transcriptional control elements in the untranslated leader of the heat shock gene *hsp22. Cell* **44:** 429–438.

Inouye, S., K. Hattori, S. Yuki, and K. Saigo. 1986. Structural variations in the *Drosophila* retrotransposon, 17.6. *Nucleic Acids Res.* **14:** 4765–4778.

Jarrell, K.A. and M. Meselson. 1991. *Drosophila* retrotransposon promoter includes an essential sequence at the initiation site and requires a downstream sequence for full activity. *Proc. Natl. Acad. Sci.* **88:** 102–104.

Javahery, R., A. Khachi, K. Lo, B. Zenzie-Gregory, and S. Smale. 1994. DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol. Cell. Biol.* **14:** 116–127.

Kadonaga, J.T. 1990. Assembly and disassembly of the *Drosophila* RNA polymerase II complex during transcription. *J. Biol. Chem.* **265:** 2624–2631.

Kerrigan, L.A., G.E. Croston, L.M. Lira, and J.T. Kadonaga. 1991. Sequence-specific transcriptional antirepression of the *Drosophila Krüppel* gene by the GAGA factor. *J. Biol. Chem.* **266:** 574–582.

Ketchum, A.S., C.T. Stewart, M. Stewart, and D.P. Kiehart. 1990. Complete sequence of the *Drosophila* nonmuscle myosin heavy-chain transcript: Conserved sequences in the myosin tail and differential splicing in the 5′ untranslated sequence. *Proc. Natl. Acad. Sci.* **87:** 6316–6320.

Kokubo, T., D.-W. Gong, S. Yamashita, M. Horikoshi, R.G. Roeder, and Y. Nakatani. 1993a. *Drosophila* 230-kD TFIID subunit, a functional homolog of the human cell cycle gene product, negatively regulates DNA binding of the TATA box-binding subunit of TFIID. *Genes & Dev.* **7:** 1033–1046.

Kokubo, T., R. Takada, S. Yamashita, D.-W. Gong, R.G. Roeder, M. Horikoshi, and Y. Nakatani. 1993b. Identification of TFIID components required for transcriptional activation by upstream stimulatory factor. *J. Biol. Chem.* **268:** 17554–17558.

Kokubo, T., D.-W. Gong, S. Yamashita, R. Takada, R.G. Roeder, M. Horikoshi, and Y. Nakatani. 1993c. Molecular cloning, expression, and characterization of the *Drosophila* 85-kilodalton TFIID subunit. *Mol. Cell. Biol.* **13:** 7859–7863.

Kokubo, T., D.-W. Gong, J.C. Wootton, M. Horikoshi, R.G. Roeder, and Y. Nakatani. 1994. Molecular cloning of Drosophila TFIID subunits. *Nature* **367:** 484–487.

Koleske, A.J. and R.A. Young. 1995. The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends Biochem. Sci.* **20:** 113–116.

Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44:** 283–292.

Lin, Y.-S., M.F. Carey, M. Ptashne, and M.R. Green. 1988. GAL4 derivatives function alone and synergistically with mammalian activators in vitro. *Cell* **54:** 659–664.

Livak, K.J. 1990. Detailed structure of the *Drosophila melanogaster Stellate* genes and their transcripts. *Genetics* **124:** 303–316.

Maldonado, E. and D. Reinberg. 1995. News on initiation and elongation of transcription by RNA polymerase II. *Curr. Opin. Cell Biol.* **7:** 352–361.

Maldonado, E., I. Ha, P. Cortes, L. Weis, and D. Reinberg. 1990. Factors involved in specific transcription by mammalian RNA polymerase II: role of transcription factors IIA, IID, and IIB during formation of a transcription-competent complex. *Mol. Cell. Biol.* **10:** 6335–6347.

McLean, C., A. Bucheton, and D.J. Finnegan. 1993. The 5′ untranslated region of the I factor, a long interspersed nuclear element-like retrotransposon of *Drosophila melanogaster*, contains an internal promoter and sequences that regulate expression. *Mol. Cell. Biol.* **13:** 1042–1050.

Mizrokhi, L.J. and A.M. Mazo. 1990. Evidence for horizontal transmission of the mobile element *jockey* between distant *Drosophila* species. *Proc. Natl. Acad. Sci.* **87:** 9216–9220.

Mizrokhi, L.J., S.G. Georgieva, and Y.V. Ilyin. 1988. *jockey*, a mobile *Drosophila* element similar to mammalian LINEs, is transcribed from the internal promoter by RNA polymerase II. *Cell* **54:** 685–691.

Mlodzik, M. and W.J. Gehring. 1987. Expression of the *caudal*

gene in the germ line of *Drosophila*: formation of an RNA and protein gradient during early embryogenesis. *Cell* **48:** 465–478.

Mlodzik, M., A. Fjose, and W.J. Gehring. 1988. Molecular structure and spatial expression of a homeobox gene from the *labial* region of the Antennapedia-complex. *EMBO J.* **7:** 2569–2578.

Moses, K., M.C. Ellis, and G.M. Rubin. 1989. The *glass* gene encodes a zinc finger protein required by *Drosophila* photoreceptor cells. *Nature* **340:** 531–536.

Nakajima, N., M. Horikoshi, and R.G. Roeder. 1988. Factors involved in transcription by mammalian RNA polymerase II: purification, genetic specificity, and TATA box-promoter interactions of TFIID. *Mol. Cell. Biol.* **8:** 4028–4040.

O'Hare, K., C. Murphy, R. Levis, and G. Rubin. 1984. DNA sequence of the *white* locus of *Drosophila melanogaster*. *J. Mol. Biol.* **180:** 437–455.

Paterson, J. and K. O'Hare. 1991. Structure and transcription of the *singed* locus of *Drosophila melanogaster*. *Genetics* **129:** 1073–1084.

Pazin, M.J., R.T. Kamakaka, and J.T. Kadonaga. 1994. ATP-dependent nucleosome reconfiguration and transcriptional activation from preassembled chromatin templates. *Science* **266:** 2007–2011.

Perkins, K.K., G.M. Dailey, and R. Tjian. 1988. In vitro analysis of the *Antennapedia* P2 promoter: identification of a new *Drosophila* transcription factor. *Genes & Dev.* **2:** 1615–1626.

Peterson, M.G., N. Tanese, B.F. Pugh, and R. Tjian. 1990. Functional domains and upstream activation properties of cloned human TATA binding protein. *Science* **248:** 1625–1630.

Pirrotta, V. 1988. Vectors for P-mediated transformation in *Drosophila*. In *Vectors: A survey of molecular cloning vectors and their uses*, (ed. R.L. Rodriguez and D.R. Denhardt) pp. 437–456. Butterworth, Boston, MA.

Poon, D., Y. Bai, A.M. Campbell, S. Bjorklund, Y.J. Kim, S. Zhou, R.D. Kornberg, and P.A. Weil. 1995. Identification and characterization of a TFIID-like multiprotein complex from *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.A* **92:** 8224–8228.

Pugh, B.F. and R. Tjian. 1992. Diverse transcriptional functions of the multisubunit eukaryotic TFIID complex. *J. Biol. Chem.* **267:** 679–682.

Purnell, B.A., P.A. Emanuel, and D.S. Gilmour. 1994. TFIID sequence recognition of the initiator and sequences farther downstream in *Drosophila* class II genes. *Genes & Dev.* **8:** 830–842.

Quan, F. and M.A. Forte. 1990. Two forms of *Drosophila melanogaster* Gsα are produced by alternate splicing involving an unusual splice site. *Mol. Cell. Biol.* **10:** 910–917.

Rubin, G.M. and A.C. Spradling. 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218:** 348–353.

Sadowski, I., J. Ma, S. Triezenberg, and M. Ptashne. 1988. GAL4-VP16 is an unusually potent transcriptional activator. *Nature* **335:** 563–564.

Sawadogo, M. and R.G. Roeder. 1985. Interactions of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* **43:** 165–175.

Schmidt, M.C., C.C. Kao, R. Pei, and A.J. Berk. 1989. Yeast TATA-box transcription factor gene. *Proc. Natl. Acad. Sci.* **86:** 7785–7789.

Schneuwly, S., A. Kuroiwa, and W.J. Gehring. 1987. Molecular analysis of the dominant homeotic *Antennapedia* phenotype. *EMBO J.* **6:** 201–206.

Segraves, W.A. and D.S. Hogness. 1990. The *E75* ecdysone-inducible gene responsible for the 75B early puff in *Drosophila*

encodes two new members of the steroid receptor superfamily. *Genes & Dev.* **4:** 204–219.

Sinclair, J.H., J.F. Burke, D. Ish-Horowicz, and J.H. Sang. 1986. Functional analysis of the transcriptional control regions of the *copia* transposable element. *EMBO J.* **5:** 2349–2354.

Singer, V.L., C.R. Wobbe, and K. Struhl. 1990. A wide variety of DNA sequences can functionally replace a yeast TATA element for transcriptional activation. *Genes & Dev.* **4:** 636–645.

Smale, S.T. 1994. Core promoter architecture for eukaryotic protein-coding genes. In *Transcription: Mechanisms and regulation*, pp. 63–80 (ed. R.C. Conaway and J.W. Conaway), Raven Press, New York, NY.

Smale, S.T. and D. Baltimore. 1989. The "initiator" as a transcription control element. *Cell* **57:** 103–113.

Soeller, W.S., S.J. Poole, and T. Kornberg. 1988. In vitro transcription of the *Drosophila engrailed* gene. *Genes & Dev.* **2:** 68–81.

Sypes, M.A. and D.S. Gilmour. 1994. Protein/DNA crosslinking of a TFIID complex reveals novel interactions downstream of the transcription start. *Nucleic Acids Res.* **22:** 807–814.

Thummel, C.S. 1989. The *Drosophila E74* promoter contains essential sequences downstream from the start site of transcription. *Genes & Dev.* **3:** 782–792.

Tyree, C.M., C.P. George, L.M. Lira-DeVito, S.L. Wampler, M.E. Dahmus, L. Zawel, and J.T. Kadonaga. 1993. Identification of a minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymerase II. *Genes & Dev.* **7:** 1254–1265.

Verrijzer, C.P., K. Yokomori, J.-L. Chen, and R. Tjian. 1994. *Drosophila* TAF$_{II}$150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA. *Science* **264:** 933–941.

Verrijzer, C.P., J.-L. Chen, K. Yokomori, and R. Tjian. 1995. Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell* **81:** 1115–1125.

Wampler, S.L., C.M. Tyree, and J.T. Kadonaga. 1990. Fractionation of the general RNA polymerase II transcription factors from *Drosophila* embryos. *J. Biol. Chem.* **265:** 21223–21231.

Weis, L. and D. Reinberg. 1992. Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. *FASEB J.* **6:** 3300–3309.

Wiley, S.R., R.J. Kraus, and J.E. Mertz. 1992. Functional binding of the "TATA" box binding component of transcription factor TFIID to the −30 region of TATA-less promoters. *Proc. Natl. Acad. Sci.* **89:** 5814–5818.

Yokomori, K., A. Admon, J.A. Goodrich, J.-L. Chen, and R. Tjian. 1993. *Drosophila* TFIIA-L is processed into two subunits that are associated with the TBP/TAF complex. *Genes & Dev.* **7:** 2235–2245.

Yokomori, K., M.P. Zeidler, J.-L. Chen, C.P. Verrijzer, M. Mlodzik, and R. Tjian. 1994. *Drosophila* TFIIA directs cooperative binding with TBP and mediates transcriptional activation. *Genes & Dev.* **8:** 2313–2323.

Zawel, L. and D. Reinberg. 1992. Advances in RNA polymerase II transcription. *Curr. Opin. Cell Biol.* **4:** 488–495.

———. 1995. Common themes in assembly and function of eukaryotic transcription complexes. *Annu. Rev. Biochem.* **64:** 533–561.

Zenzie-Gregory, B., A. Khachi, I.P. Garraway, and S.T. Smale. 1993. Mechanism of initiator-mediated transcription: evidence for a functional interaction between the TATA-binding protein and DNA in the absence of a specific recognition sequence. *Mol. Cell. Biol.* **13:** 3841–3849.

Zhou, Q., P.M. Lieberman, T.G. Boyer, and A.J. Berk. 1992. Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes & Dev.* **6:** 1964–1974.

Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters.

T W Burke and J T Kadonaga

References	This article cites 82 articles, 41 of which can be accessed free at: **http://genesdev.cshlp.org/content/10/6/711.full.html#ref-list-1**	
License		
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**	