TATA-binding protein-associated factor(s) in TFIID function through the initiator to direct basal transcription from a TATA-less class II promoter

Ernest Martinez, Cheng-Ming Chiang,
Hui Ge and Robert G. Roeder¹

Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA

¹Corresponding author

Communicated by G. Blobel

The RNA polymerase II (Pol II) basal transcription factor TFIID is composed of the TATA box-binding protein (TBP) and several TBP-associated factors (TAFs). TBP is required for Pol II transcription from TATA-containing and TATA-less promoters. TATA-less promoters of mRNA-encoding genes often contain an initiator element at the transcription start site that is sufficient to direct accurate Pol II transcription. Here we address the mechanisms of functional TBP recruitment to the TATA-less initiator-dependent promoter of the mouse terminal deoxynucleotidyl transferase (TdT) gene. We show that the natural TATA-less TdT initiator region is sufficient to promote low levels of specific transcription *in vitro* and to direct the assembly of a stable preinitiation complex. In contrast to what is observed for several other promoters lacking a consensus TATA element, the TATA-binding activity of TBP is not required for the functional recruitment of TFIID to the natural TATA-less TdT and β -polymerase promoters. Moreover, a comparison of TBP and highly purified epitope-tagged TFIID reveals that one or several TAFs function independently of distal regulatory elements to mediate initiator-directed (basal) transcription from the natural TATA-less TdT core promoter in crude nuclear extracts. Furthermore, by using a transcription system reconstituted with purified components, we present the first evidence for a basal transcription function of TAFs through the TdT initiator element. Altogether, our results suggest an alternative pathway for TFIID recruitment to initiator-dependent TATA-less class II promoters in which TAF(s) recruit TBP by interacting either directly or indirectly with the initiator region.

Key words: initiator/TAFs/ TATA-less promoter/TFIID/transcription

Introduction

Specific transcription initiation at class II gene promoters requires the assembly of a precisely positioned RNA polymerase II (Pol II)-containing preinitiation complex over the core promoter DNA region. Most mammalian class II core promoters typically contain a TATA box ~25–30 bp upstream of the transcription initiation site and/or an initiator element located at the transcription start site, each of which can independently specify the location and direction of

transcription initiation (Breathnach and Chambon, 1981; O'Shea-Greenfield and Smale, 1992). The process of functional transcription preinitiation complex formation on TATA-containing class II core promoters involves the ordered assembly of at least seven basal transcription initiation factors including TFIIA, -B, -D, -E, -F, -G/J and -H, in addition to Pol II. TFIID is the only one of these factors to possess sequence-specific DNA-binding activity and initiates preinitiation complex assembly through recognition of the TATA box (reviewed in Roeder, 1991; Zawel and Reinberg, 1992). In higher eukaryotes, native TFIID is a large multisubunit complex composed of a small TATA-binding protein (TBP) and several tightly associated proteins commonly referred to as TBP-associated factors (TAFs) (Dymlacht *et al.*, 1991; Pugh and Tjian, 1991; Tanese *et al.*, 1991; Takada *et al.*, 1992; Zhou *et al.*, 1992; Chiang *et al.*, 1993). Importantly, in purified reconstituted systems TBP can function in the absence of TAFs to nucleate the assembly of a complete functional preinitiation complex promoting specific basal transcription from TATA-containing core promoters (Buratowski *et al.*, 1988; Hoffmann *et al.*, 1990; Peterson *et al.*, 1990; Pugh and Tjian, 1990; Cortes *et al.*, 1992; Parvin and Sharp, 1993; and references therein). However, effective responses to activators in such systems require the TAFs in TFIID (reviewed in Chiang *et al.*, 1993), as well as distinct components in the upstream factor stimulatory activity (USA) cofactor fraction (Meisterernst *et al.*, 1991).

Recently, TBP has been shown to be required for *in vitro* transcription of class II genes that do not contain a consensus TATA box (Pugh and Tjian, 1990; Smale *et al.*, 1990; Carcamo *et al.*, 1991; Pugh and Tjian, 1991; Yoganathan *et al.*, 1992; Zhou *et al.*, 1992; Bernués *et al.*, 1993), as well as for the TATA-less ribosomal (class I) genes transcribed by RNA polymerase I (Pol I) (Comai *et al.*, 1992) and for TATA-containing and TATA-less class III genes transcribed by RNA polymerase III (Pol III) (Lobo *et al.*, 1991; Margottin *et al.*, 1991; Simmen *et al.*, 1991, 1992; Huet and Sentenac, 1992; Kassavetis *et al.*, 1992; Lobo *et al.*, 1992; Taggart *et al.*, 1992; White and Jackson, 1992; White *et al.*, 1992). These results were further confirmed *in vivo* by mutagenesis of the TBP gene in yeast (Cormack and Struhl, 1992; Schultz *et al.*, 1992) and underscore the importance of TBP not only for TATA-dependent but also, and more interestingly, for TATA-independent transcription by all three nuclear RNA polymerases.

How is TBP recruited to TATA-less promoters? In the case of class I genes, TBP is part of the multisubunit selectivity factor SL1 which is recruited to the core ribosomal RNA promoter by protein–protein interactions with the upstream promoter-bound activator UBF and by species-specific TAF(s)–core promoter DNA interactions (Learned *et al.*, 1986; Bell *et al.*, 1990; Comai *et al.*, 1992). In the

case of TATA-less class III genes, TBP is part of the TFIIB multisubunit complex which is tethered to the core promoter by protein–protein interactions with TFIIC bound to the intragenic core promoter elements (Huet and Sentenac, 1992; Kassavetis *et al.*, 1992; Simmen *et al.*, 1992; Taggart *et al.*, 1992; White and Jackson, 1992; Chiang *et al.*, 1993). In contrast, the mechanisms of TBP recruitment to TATA-less mRNA-encoding class II gene promoters are still unclear and several models have been proposed. For instance, Wiley *et al.* (1992) observed TBP binding *in vitro* to the –30 region of several natural promoters lacking a consensus TATA box and proposed that as for TATA-containing promoters, TBP binding to the –30 DNA region is also necessary for TATA-less promoter function. Similarly, Zenzie-Gregory *et al.* (1993a) reported that different synthetic DNA sequences positioned at –30 from the mouse terminal deoxynucleotidyl transferase (TdT) initiator element differentially affect Sp1-activated transcription *in vitro*, and suggested that TBP must bind to the –30 region of TATA-less promoters regardless of the sequences in that region. In contrast, Pugh and Tjian (1991) reported that Sp1-dependent activation of a synthetic initiator-containing promoter lacking a consensus TATA element involved TBP recruitment to the promoter solely by the interaction of upstream promoter-bound Sp1 molecules with a heat-labile tethering TAF(s) in the TFIID complex. Another mechanism was proposed by Carcamo *et al.* (1991) who reported that purified Pol II can weakly recognize the adenovirus major late promoter initiator and directs the cooperative recruitment of TBP and the other basal initiation factors in the absence of a consensus TATA element. Finally, consistent with a possible diversity in initiator elements among various promoters (Weis and Reinberg, 1992), different factors including HIP1 (Means and Farnham, 1990; Means *et al.*, 1992), YY1 (Seto *et al.*, 1991; Shi *et al.*, 1991), TFII-I (Roy *et al.*, 1991) and the previously cloned 43 kDa helix–loop–helix activator USF (Gregor *et al.*, 1990; Roy *et al.*, 1991; Du *et al.*, 1993; and references therein) have been identified recently as sequence-specific initiator-binding proteins. Three of these have been shown to stimulate transcription through an initiator element. YY1 stimulates basal transcription *in vitro* from the adeno-associated virus (AAV) P5 initiator (Seto *et al.*, 1991). TFII-I is required for initiator-dependent basal transcription from the TATA-containing adenovirus major late (AdML) core promoter in the absence of TFIIA in a system reconstituted with purified basal transcription factors and Pol II (Roy *et al.*, 1991). USF has been shown to stimulate transcription *in vivo* through the initiator sequences of AdML and HIV-1 promoters (Du *et al.*, 1993). Altogether, these observations led to the proposal that initiator-binding factors can promote basal TATA-less transcription by facilitating and/or stabilizing TBP binding to the core promoter (Roeder, 1991; Weis and Reinberg, 1992).

In essence, the mechanisms proposed previously for TBP recruitment to different class II promoters that lack a consensus TATA box include (i) the direct binding of TBP to low-affinity non-consensus binding sites in the –30 promoter region, (ii) the tethering of TBP by upstream promoter-bound activators interacting with TAF(s), and (iii) interaction of TBP with sequence-specific initiator-binding proteins or Pol II. In contrast, here we present evidence that functional TFIID recruitment to a natural TATA-less core

promoter requires neither its TATA-binding activity nor the action of upstream promoter-bound activators. Indeed, we show that in contrast to TATA-dependent transcription, initiator-directed basal transcription from the natural TATA-less TdT core promoter is resistant to specific competition with a consensus TATA oligonucleotide. This specific competition assay allows us to distinguish ‘true’ TATA-less promoters like the natural TdT and β -polymerase (β -Pol) promoters from promoters that lack a consensus TATA element but still have functional low-affinity TBP-binding sites, such as those that have been used in most, if not all, of the previous studies (see above) on supposedly TATA-less promoters (Pugh and Tjian, 1990, 1991; Seto *et al.*, 1991; Wiley *et al.*, 1992; Yoganathan *et al.*, 1992; Zhou *et al.*, 1992; Zenzie-Gregory *et al.*, 1993a). Consistent with these results, we further demonstrate that although TBP can direct basal transcription *in vitro* from TATA-containing promoters, it cannot promote specific initiator-dependent basal transcription from the ‘true’ TATA-less TdT core promoter unless TFIID-specific TAFs are present. This suggests that direct TBP–DNA interactions do not make primary contributions to TBP recruitment to the TdT promoter; however, as observed for TATA-containing promoters, we show that the TATA-less TdT core promoter also directs the formation of a stable Sarkosyl- and heparin-resistant preinitiation complex. Moreover, these results demonstrate, for the first time, that one or several components tightly associated with TBP in the TFIID complex (i.e. TAFs) are required independently of upstream activators for initiator-directed basal transcription from a ‘true’ TATA-less core promoter. Furthermore, by using a transcription system reconstituted with purified components, we demonstrate a basal transcription function of TAFs through the wild-type, but not a mutant, TdT initiator element. Therefore, our results suggest a possible alternative TFIID recruitment pathway in which TAF(s) can tether TBP to a class II core promoter by either direct or indirect interactions with the initiator region.

Results

Efficient basal transcription from the TdT promoter is directed by the initiator and its natural downstream region

To study basal initiator-dependent TATA-less transcription independently of the action of upstream activators, we analyzed specific transcription from the natural mouse TdT promoter in a HeLa cell nuclear extract. This promoter lacks a TATA box but contains a functional initiator element and its expression *in vivo* is lymphoid-specific. Indeed, previous *in vivo* transfection and *in vitro* transcription analyses of TdT promoter deletion mutants showed that TdT-expressing lymphoid cell lines direct activated transcription through lymphoid-specific sequences located upstream of –41. By contrast, TdT-lacking cell lines like HeLa cells only support basal transcription directed by the initiator region (Smale and Baltimore, 1989; Lo *et al.*, 1991; E.Martinez and R.G.Roeder, unpublished results). To confirm that TdT upstream promoter sequences do not activate initiator-mediated transcription under our *in vitro* conditions, we compared the levels of specific transcription directed by the wild-type TdT promoter (Figure 1A, lane 1) with a truncated TdT promoter lacking sequences upstream of –41 (lane 2).

Fig. 1. Basal initiator-directed *in vitro* transcription from the TATA-less TdT promoter is independent of upstream promoter sequences. (A) Structural features of TdT promoter constructs. The initiator (INR box, positions -6 to +11; Smale *et al.*, 1990) and its +33/+59 downstream region (DR) are indicated in the top scheme. Constructs 1-4 are respectively: pTdT(-1300/+59), pTdT(-41/+59), pTdT(-1300/+33) and pTdT(-41/+33). (B) Specific *in vitro* transcription activity of the different constructs presented in (A) and normalized to the activity of construct 1 (100%) in a HeLa nuclear extract. Correctly initiated Pol II-specific TdT transcripts (sensitive to 1 μ g/ml α -amanitin, data not shown) were analyzed as described in Materials and methods and quantitated by counting the radioactivity present in cut gel slices. Data represent mean values from at least two transcription experiments. Almost identical relative transcription activities were obtained with either 5 or 10 μ l of nuclear extract, although the absolute level of TdT transcription increases ~4-fold with 10 μ l of nuclear extract (data not shown). (C) TdT core promoter constructs containing either the wild-type TATA-less -25 to -31 DNA sequence (TATA⁻, constructs 2 and 4) or a consensus TATA box in its place (TATA⁺, constructs 5 and 6) and containing (+59, constructs 2 and 5) or lacking (+33, constructs 4 and 6) the +33/+59 DR sequence. (D) Transcription with TATA⁻ and TATA⁺ constructs shown in (C). Transcriptions were carried out, as described in Materials and methods, by mixing 80 fmol of each +59 and +33 template with 5 μ l of nuclear extract. Arrows indicate the products of correctly initiated transcripts. Lane 3 is a 60 times shorter exposure of lane 2. The amount of correctly initiated transcripts was quantitated as above. The amount of transcription obtained with the +59 construct versus the +33 construct (+59/+33) is 3.5-fold for TATA⁻ templates and 1.7-fold for TATA⁺ templates, and was independent of the amount of extract used for transcription (data not shown).

Correctly initiated Pol II transcripts were analyzed by primer-extension and quantitated as described in the legend to Figure 1. Comparison of lanes 1 and 2 in Figure 1B shows, as expected, that TdT upstream promoter sequences do not activate initiator-directed transcription in a HeLa nuclear extract. Consistent with a previous observation (Smale and Baltimore, 1989), deletion of the downstream region (DR) between +33 and +59 (Figure 1A) decreases transcription 3- to 4-fold (Figure 1B, lane 1 versus 3). The additional deletion of the upstream sequences in this DR-deleted TdT construct (Figure 1A, lane 4) does not further reduce transcription (Figure 1B, compare lane 3 with 4), confirming the absence of functional TdT upstream promoter-binding activators in the HeLa nuclear extract. Similarly, the additional deletions of sequences between +11 and +33 and between -41 and -6 also do not further decrease basal transcription, whereas deletion of the initiator element (sequences -6 to +11) abolishes specific transcription (Smale and Baltimore, 1989; Zenzie-Gregory *et al.*, 1993a).

We thus conclude that the DR sequence is the only initiator-flanking region that significantly stimulates basal TdT transcription. Since deletion of the DR sequence does not change RNA stability in the HeLa nuclear extract (data not shown), the function of this downstream sequence was further addressed. To test whether the DR sequence functions as a typical activator element able to stimulate transcription in the presence of a TATA box, we replaced the -25 to -31 natural TdT sequence with a consensus TATA element in promoter constructs with and without the DR sequence (Figure 1C). The DR activity was then compared in the presence of the consensus TATA box (Figure 1C, TATA⁺, constructs 5 versus 6) and with the wild-type TATA-less sequence (Figure 1C, TATA⁻, constructs 2 versus 4). The results presented in Figure 1D show that replacing the natural -30 TdT DNA region by the consensus TATA sequence strongly stimulates basal TdT transcription (compare lane 1 with 2). More interestingly, the moderate stimulatory activity of the DR sequence observed in the wild-type TATA-

less TdT promoter context (lane 1, compare +59 with +33) is reduced after insertion of the TATA box (lanes 2 and 3, compare +59 with +33). This suggests that the TdT DR does not function as a typical activator element but rather that it might contribute to TFIID recruitment by a mechanism that is not required when a consensus TATA box is present.

From these results we conclude that specific transcription from the natural TATA-less TdT promoter in a HeLa nuclear extract does not require the function of typical distal promoter-bound activators; instead, it reflects basal promoter activity directed by the TdT core promoter which consists minimally of the initiator element (sequences -6 to $+11$) or, for maximal activity, of the initiator and its natural downstream region (sequences -6 to $+59$) (Smale and Baltimore, 1989; Smale *et al.*, 1990). Importantly, the previously described synthetic promoter construct pSp1 42 INR (plasmid VI in Smale *et al.*, 1990), and various derivatives that have been used in studies to date to address the mechanisms of TATA-independent initiator-directed transcription (Pugh and Tjian, 1990, 1991; Seto *et al.*, 1991; Zenzie-Gregory *et al.*, 1992; Zhou *et al.*, 1992; Jiang *et al.*, 1993), actually contain a weak but functional TATA-like plasmid DNA sequence at -30 upstream of the initiator (see below and Zenzie-Gregory *et al.*, 1993a). Therefore, to prevent non-consensus TBP-binding sites present in flanking plasmid sequences from interfering in our TATA-less transcription assays, our following analyses of basal (core) TATA-less initiator-dependent transcription have been conducted with TdT promoter constructs that contain natural flanking DNA sequences extending at least 41 bp upstream and either 59 or 33 bp downstream (for an active or inactive DR, respectively) of the initiation site.

The TATA-binding activity of TBP is not required for the functional recruitment of TFIID to 'true' TATA-less promoters

Previous studies have demonstrated that TBP can bind *in vitro* to a variety of non-consensus TATA box DNA sequences (Hahn *et al.*, 1989), as well as to sequences present in the -30 DNA region of several promoters lacking a recognizable TATA element (Wiley *et al.*, 1992). In addition, several different DNA sequences can functionally replace a TATA element *in vivo* and *in vitro* (Singer *et al.*, 1990; Zenzie-Gregory *et al.*, 1993a). These observations led to the proposal that the binding of TBP to non-TATA sequences is required for TATA-less initiator-directed transcription (Zenzie-Gregory *et al.*, 1993a) and may contribute to TFIID recruitment to TATA-less promoters (Wiley *et al.*, 1992). To address the mechanism of TBP recruitment to the natural TATA-less TdT core promoter, we analyzed the role of TBP–DNA interactions in TdT promoter function by testing whether the TATA DNA-binding activity of TBP is required for specific TATA-less initiator-directed transcription *in vitro*. This was tested by preincubating HeLa nuclear extract with an excess of either a non-specific control or a consensus AdML TATA-containing oligonucleotide to block the TATA DNA-binding function of TFIID before initiating transcription by the addition of nucleotides and the template promoters. Figure 2A shows that preincubation of the nuclear extract with the TATA oligonucleotide specifically and strongly inhibits transcription from a synthetic E1b TATA-dependent

Fig. 2. Blocking the TATA-binding function of TFIID with a TATA oligonucleotide inhibits transcription from TATA box-containing promoters and from a synthetic promoter lacking a consensus TATA element, but not from the TATA-less TdT and β -polymerase promoters. (A) Selective inhibition of TATA-dependent transcription, but not TATA-less TdT initiator-directed transcription, by preincubation of a HeLa nuclear extract with a TATA oligonucleotide. Specific transcription from pTdT($-41/+59$), pE1b–CAT and pAdML($-257/+33$)–CAT (arrows labeled TdT+59, E1b and AdML, respectively) was analyzed in a nuclear extract preincubated without (lane 1) and with a 150-fold molar excess over the template of either a TATA oligonucleotide (lane 2) or a control non-specific oligonucleotide (NS, lane 3), as described in Materials and methods. (B) Addition of recombinant TBP restores TATA-dependent, but not 'true' TATA-less promoter transcription in TATA oligonucleotide-blocked nuclear extracts. Specific transcription from pTdT($-1300/+59$), pHsp70($-33/+99$)–CAT, p β P10 and pSp1 42 INR (arrows labeled TdT+59, Hsp70, β -Pol and Sp1 42 INR, respectively) was analyzed as described in (A) and without (lanes 1–3) or with the addition of 10 ng of recombinant f:TBP (lane 4), as described in Materials and methods. Note that the strength of the signals in the different panels does not correspond to relative promoter strength of the templates used because different exposures are presented.

promoter and from the natural TATA-containing AdML promoter, but not from the TATA-less TdT core promoter (compare lane 2 with 1), whereas preincubation with a control oligonucleotide (NS) has no effect (lane 3 versus 1). An independent experiment presented in Figure 2B further shows that the TATA oligonucleotide inhibits specific transcription from the human TATA-containing Hsp70 core promoter but again does not affect transcription from the TdT promoter (compare lane 2 with 3 for TdT+59 and Hsp70).

To confirm that the TATA-binding subunit of TFIID is the factor blocked specifically by the TATA oligonucleotide

and to test further whether TBP–DNA interactions are limiting in the nuclear extract for transcription from the TdT promoter, purified epitope-tagged recombinant TBP (f:TBP; see Figure 4, lane 3) was preincubated with the template promoter before combining with the nuclear extract preblocked with the TATA oligonucleotide (see Materials and methods). As shown in Figure 2B, ectopic TBP overcomes the inhibitory effect of the competitor TATA oligonucleotide on TATA-dependent transcription from the Hsp70 core promoter (compare lane 3 with 4 for Hsp70). Consistent with the observation that preincubation of the nuclear extract with the TATA oligonucleotide does not inhibit transcription from the TdT core promoter, preincubation of the template with TBP prior to addition to the nuclear extract does not stimulate, but instead inhibits, specific transcription (compare lane 3 with 4 for TdT + 59). Neither inhibition nor stimulation of TdT transcription is observed when TBP and the TdT template are added separately to a nuclear extract, whereas TATA-dependent Hsp70 promoter transcription is also stimulated under these conditions (data not shown). Taken together, these results suggest that the TATA DNA-binding activity of TBP is not required for specific initiator-directed transcription from the TdT promoter.

By using the same assay, we also tested a possible requirement of the TATA-binding activity of TBP for transcription from the natural TATA-less human β -Pol promoter (plasmid p β P10 in Widen *et al.*, 1988) and the synthetic TdT initiator-dependent pSp1 42 INR construct that lacks a consensus TATA box (plasmid VI in Smale *et al.*, 1990). As observed for the TdT promoter, transcription from the natural TATA-less β -Pol promoter is not specifically inhibited by preincubation of the nuclear extract with the TATA oligonucleotide (Figure 2B, β -Pol, compare lane 2 with 3) and preincubation of TBP with the promoter is again inhibitory (β -Pol, lane 4). In contrast, transcription from the synthetic pSp1 42 INR construct is specifically inhibited by the TATA oligonucleotide (Sp1 42 INR, compare lane 2 with 3). Moreover, preincubation of the template with ectopic TBP does not restore the original level of Sp1-activated transcription (Sp1 42 INR, compare lane 1 with 4), nor does it inhibit transcription from this promoter (compare lane 3 with 4). The latter result is consistent (i) with the binding of the TATA oligonucleotide to the TBP subunit in the TFIID complex, thus preventing its interaction with a weak but functional AT-rich plasmid sequence (i.e. ATATCAT) located at –30 in this synthetic promoter (see also Zenzie-Gregory *et al.*, 1993a,b), and (ii) with earlier studies demonstrating the requirement for TBP-associated coactivators present in the TFIID complex for Sp1-activated transcription from this template (Pugh and Tjian, 1990; Zhou *et al.*, 1992). The differential effect of ectopic TBP on the TATA oligonucleotide-inhibited transcription from the Hsp70 and Sp1 42 INR templates simply reflects the fact that the Hsp70 activity is solely TBP-dependent core promoter activity, whereas Sp1 42 INR activity in the nuclear extract is mostly the result of Sp1-dependent activated transcription. Indeed, TBP promotes specific basal transcription from the pSp1 42 INR template in a purified reconstituted transcription system in the absence of Sp1 activator (data not shown).

In summary, these experiments demonstrate that in

contrast to TATA-containing promoters and some presumptive TATA-less promoters having functional low-affinity TBP binding sites (like pSp1 42 INR; see also Zenzie-Gregory *et al.*, 1993a), the natural TdT and β -Pol promoters do not require the TATA DNA-binding activity of TBP for the functional recruitment of TFIID (see also below). Based on this criterion, the TdT and β -Pol promoters can be considered as ‘true’ TATA-less promoters, whereas at least some of the previously analyzed synthetic and natural promoters that lack consensus TATA elements (Wiley *et al.*, 1992; Yoganathan *et al.*, 1992; Zenzie-Gregory *et al.*, 1993a) should be considered as weak TATA-containing promoters.

Sarkosyl- and heparin-resistant preinitiation complex formation on the TATA-less TdT core promoter

The anionic detergent Sarkosyl has been used previously to separate and characterize functional steps in transcription initiation by Pol II from the AdML promoter in different *in vitro* transcription systems (Hawley and Roeder, 1985, 1987; Reinberg *et al.*, 1987). Sarkosyl added to a nuclear extract at concentrations >0.025% inhibits the formation of a complete functional preinitiation complex on the consensus TATA-containing AdML promoter, whereas concentrations >0.08% are required to prevent productive transcription initiation once a complete preinitiation complex is formed (Hawley and Roeder, 1987). Interestingly, both TFIID and TFIIA are required during the preincubation step to form a Sarkosyl-resistant preinitiation complex on the AdML promoter, underscoring the importance of TBP–TATA box interactions in the assembly of a stable preinitiation complex on a TATA-containing promoter (Reinberg *et al.*, 1987). Similarly, the synthetic Sp1-activated weak TATA-containing promoter construct pSp1 42 INR (see above and Zenzie-Gregory *et al.*, 1993a,b) has also been shown to direct the formation of a Sarkosyl-resistant preinitiation complex (Zenzie-Gregory *et al.*, 1992). To test whether a natural ‘true’ TATA-less core promoter can direct the formation of a stable Sarkosyl-resistant preinitiation complex independently of upstream activators, we analyzed the Sarkosyl sensitivity of specific transcription from the natural TATA-less TdT core promoter in a HeLa nuclear extract. As outlined in Figure 3C, our transcription reactions were carried out by preincubating template promoters with the nuclear extract for 50 min at 30°C (preinitiation complex formation). After the preincubation, transcription was initiated by adding nucleotides (NTPs) and transcript elongation was then allowed to proceed for 60 min at 30°C. Increasing amounts of Sarkosyl were added either before the preincubation step (T_0) or after preincubation but 30 s before the addition of nucleotides (T_{50}). To inhibit reinitiation, the final Sarkosyl concentration was adjusted to 0.08% 30 s after the addition of nucleotides (Hawley and Roeder, 1987). Figure 3A shows transcription from the TATA-containing human Hsp70 and the TATA-less TdT core promoters when different concentrations of Sarkosyl are added to transcription reactions either before (T_0 , lanes 1–6) or after (T_{50} , lanes 7–12) the 50 min preincubation step. The transcriptional activity of the Hsp70 and TdT core promoters in response to the addition of increasing amounts of Sarkosyl at T_0 is very similar (lanes 1–6), showing a sharp inhibition of transcription at 0.03% and higher

Fig. 3. Formation of a functional Sarkosyl- and heparin-resistant preinitiation complex on the TATA-less TdT promoter. (A) Sarkosyl sensitivity of TATA-dependent Hsp70 and TATA-less initiator-dependent TdT transcription before and after preinitiation complex formation. Transcription reactions were carried out with 80 fmol of pTdT(-41/+59) and 40 fmol of pHsp70(-33/+99)-CAT in 10 μ l of nuclear extract (see Materials and methods), as described in the text and summarized below in part (C). Increasing amounts of Sarkosyl were added either at time zero (T₀, lanes 1–6) or at time 50 min after preinitiation complex formation (T₅₀, lanes 7–12) and before the addition of nucleotides. Arrows indicate the position of correctly initiated transcripts. (B) Heparin-resistant TATA-less initiator-directed transcription from the TdT promoter is independent of the DR sequence. Transcription reactions were carried out as described in the text [see also (C) and Materials and methods] and contained 5 μ l of nuclear extract and 80 fmol each of pTdT(-1300/+59) and pTdT(-1300/+33) (TdT+59 and TdT+33, respectively). No heparin (lane 1) or 10 μ g/ml heparin was added at time zero (T₀, lane 2) or after preinitiation complex formation at time 50 min (T₅₀, lane 3). (C) Diagram of the experiments presented in (A) and (B). Sarkosyl or heparin was added to the reaction containing the nuclear extract (NE), the templates and all the other components except the nucleotides (NTPs), either at time zero (T₀) or after the 50 min preincubation at 30°C (T₅₀), but 30 s (Sarkosyl experiment) or 2 min (heparin experiment) before the addition of NTPs. For the experiment in (A), additional Sarkosyl was added 30 s after the NTPs to adjust the final concentration of Sarkosyl in all the reactions to 0.08% and to inhibit reinitiation (Hawley and Roeder, 1987). After the addition of NTPs, reactions were further incubated for 60 min at 30°C, stopped (T₁₁₀) and processed as described in Materials and methods.

Sarkosyl concentrations (lanes 5 and 6). The same sensitivity to Sarkosyl was observed with a TdT template lacking a functional DR sequence (data not shown). Interestingly, at T₅₀ after the preincubation step (lanes 7–12), transcription from the TATA-less TdT core promoter is resistant to a Sarkosyl concentration (0.04%), which when added at T₀ abolishes transcription from both TdT and Hsp70 promoters (compare lane 6 with 11).

Importantly, the preincubation-dependent Sarkosyl-resistant transcription (Figure 3A) results mainly from the assembly of stable preinitiation complexes on both TATA-containing (Hsp70) and TATA-less (TdT) core promoters, and not from the formation of initiated transcription complexes during the preincubation step. This is shown by the fact that 0.08% Sarkosyl inhibits transcription, but only when added 30 s before, rather than 30 s after, nucleotide addition (compare for instance lane 1 with 0.08% Sarkosyl added 30 s after NTPs and lane 12 with 0.08% Sarkosyl added 30 s before NTPs). A similar challenge assay with heparin (Figure 3C) was employed to confirm further the assembly of a stable preinitiation complex on the TATA-less TdT core promoter and to test for the contribution of

the DR sequence; indeed, like Sarkosyl, heparin allows the separation of the preinitiation step from the initiation/elongation steps (Reinberg and Roeder, 1987). Transcription from a DR-containing and a DR-lacking TdT promoter (TdT+59 and TdT+33, respectively, in Figure 3B) was analyzed in the absence and presence of heparin (10 μ g/ml) added either before the preincubation step (at T₀) or after the preincubation period but 2 min before the addition of nucleotides (at T₅₀). As shown in Figure 3B, transcription from both TdT promoter constructs is strongly inhibited by heparin added before the preincubation step (T₀, lane 2). In contrast, after the preincubation period, transcription from both templates becomes significantly resistant to heparin (compare lane 2 with 3). In addition, the relative transcription activity of the DR sequence-containing versus the DR-lacking TdT promoter is the same (~3- to 4-fold) in the absence and in the presence of heparin (compare lane 1 with 3), showing that the DR sequence is not required for the formation of a heparin-resistant preinitiation complex on the TdT promoter. Taken together, these data demonstrate that the initiator region is sufficient, in the absence of DR and distal activator elements, to direct the assembly in a HeLa

nuclear extract of a stable (i.e. Sarkosyl- and heparin-resistant) functional preinitiation complex on the 'true' TATA-less TdT core promoter.

TBP-associated factor(s) in the TFIID complex are required for TATA-independent initiator-directed basal transcription from the TdT core promoter

The results of the TATA oligonucleotide competition assay (Figure 2) suggest that direct TBP–DNA interactions do not contribute to functional TBP recruitment to the TATA-less TdT core promoter, but rather that protein–protein interactions may be involved. In a previous study, Pugh and Tjian (1991) reported that a heat-labile TAF(s) is necessary for Sp1-activated transcription from a synthetic TdT initiator-containing promoter lacking a consensus TATA element (i.e. a derivative of pSp1 42 INR). Since basal initiator-dependent transcription in the absence of Sp1 could not be detected under the conditions used, it was proposed that TAFs recruit TBP solely by interacting with the upstream promoter-bound Sp1 activators. However, our results (Figure 1) and previous observations (Smale and Baltimore, 1989; Smale *et al.*, 1990) have demonstrated that the initiator region, from either TdT or the AdML promoter, can direct basal transcription in the absence of both a TATA box and distal activator DNA elements. We thus decided to test the alternative possibility that one or several TAF(s) may actually function through the initiator region of the TATA-less TdT core promoter independently of distal promoter-bound activators. This was analyzed by adding either purified bacterially expressed recombinant Flag epitope-tagged TBP (f:TBP; Figure 4, lane 3) or a highly purified native Flag epitope-tagged TFIID complex (f:TFIID) composed of f:TBP and 13 major tightly-associated TAFs (Figure 4, lane 2; Chiang *et al.*, 1993) to a HeLa nuclear extract depleted of TFIID activity by a mild heat treatment that leaves all the other basal factors relatively unaffected (see Materials and methods; Nakajima *et al.*, 1988). Specific basal transcription from the TATA-containing Hsp70 core promoter and from two TdT promoter constructs having (TdT+59) and lacking (TdT+33) the DR sequence, was tested in parallel (Figure 5A). Heat inactivation of TFIID in the nuclear extract inhibits transcription from the TATA-containing Hsp70 and TATA-less TdT promoters (compare lane 1 with 2). Complementation of the heat-treated nuclear extract with either f:TFIID or f:TBP restores basal transcription from the TATA-containing Hsp70 promoter (compare lanes 3–5 for Hsp70). In contrast, only f:TFIID, but not f:TBP, can promote basal transcription from the TATA-less TdT promoter, independent of the presence or absence of the DR sequence (compare lanes 3–5 for TdT+59 and TdT+33). Moreover, the addition of excess f:TBP (40 ng) is not sufficient to promote specific transcription from the TdT promoter (data not shown).

To test whether specific TAFs present only in the purified f:TFIID complex are required for initiator-directed transcription from the TdT promoter, we analyzed the function of another highly purified Flag epitope-tagged TBP-containing complex f:TFIIB₂, composed of f:TBP and four specific TAFs (Figure 4, lane 1; Chiang *et al.*, 1993). TFIIB₂ is part of the Pol III transcription factor TFIIB and has been shown recently to be required for TATA-less Pol III transcription *in vitro* (Chiang *et al.*, 1993). The experiment presented in Figure 5B shows that under

Fig. 4. Polypeptide composition of highly purified Flag epitope-tagged TBP–TAFs complexes f:TFIIB₂ and f:TFIID and bacterially expressed f:TBP on silver-stained SDS gels. f:TBP–TAFs complexes were purified from nuclear extracts of the 3-10 TBP epitope-tagged HeLa cell line by using 0.3 and 0.85 M KCl phosphocellulose P11 fractions to immunoaffinity purify f:TFIIB₂ and f:TFIID, respectively, on an anti-Flag M2 monoclonal antibody-resin, as described previously (Chiang *et al.*, 1993; see Materials and methods). Lanes 1 and 2 show f:TFIIB₂ and f:TFIID preparations, respectively. The molecular weights of the major specific TAFs are indicated. Dashed lines in lane 2 indicate minor TFIID-specific polypeptides and the star in lanes 1 and 2 indicates a non-specific polypeptide (see also Chiang *et al.*, 1993). Lane 3 shows immunoaffinity purified recombinant f:TBP expressed in bacteria (Chiang and Roeder, 1993).

conditions in which f:TFIID can promote transcription from the TdT promoter (lane 4), an equivalent amount of f:TFIIB₂ (in TBP units, see legend to Figure 5) cannot promote Pol II transcription either from the TATA-less TdT promoter (lane 3) or from the TATA-containing AdML promoter (data not shown).

Interestingly, an amount of f:TFIID that restores almost completely the Hsp70 transcriptional activity observed with the nuclear extract before heat treatment (Figure 5A upper panel, lane 3) only restores a small fraction of the original TdT promoter activity (two lower panels, lane 3). This is also observed when using a partially purified natural TFIID preparation and when using an anti-TBP immuno-depleted nuclear extract instead of a heat-treated nuclear extract (data not shown). This suggests the presence of a heat-sensitive factor (only partially inactivated after heat treatment of the nuclear extract) that is able to associate with TFIID and is necessary for TATA-less initiator-directed transcription, but not for TATA-directed transcription (E.Martinez and

Fig. 5. A highly purified TFIID complex but not TBP can promote basal initiator-directed transcription from the TATA-less TdT promoter. (A) Comparison of f:TFIID and f:TBP in supporting basal TATA-dependent and initiator-dependent transcription. Specific transcription from pHsp70(−33/+99)–CAT, pTdT(−41/+59) and pTdT(−1300/+33) (arrows labeled Hsp70, TdT+59 and TdT+33, respectively) was analyzed in a nuclear extract before (lane 1, NE) and after heat inactivation (lanes 2–5, HTNE). BSA alone (2.5 μ g, lane 2) and either f:TFIID (5 TBP U with BSA, lane 3) or f:TBP (5 and 10 TBP U with BSA, lanes 4 and 5, respectively) were added to reactions containing heat-treated nuclear extract (HTNE). One TBP unit corresponds to $\sim$ 1 ng of f:TBP, as determined by Western blot and silver-staining (data not shown). (B) f:TFIID₂ does not promote basal transcription from the TdT promoter. Specific transcription from pTdT(−41/+59) (arrow labeled TdT+59) was analyzed in a nuclear extract before (lane 1) and after heat inactivation (lanes 2–4). BSA alone (0.75 μ g, lane 2) and either f:TFIID₂ (2 TBP U with BSA, lane 3) or f:TFIID (2 TBP U with BSA, lane 4) were added to the reactions with heat-treated nuclear extract (HTNE).

R.G.Roeder, unpublished results). These experiments show, for the first time, that one or several TAFs in the TFIID complex are required for initiator-directed basal transcription from the TATA-less TdT core promoter and that this occurs independently of the DR region and of upstream promoter-bound activators.

TBP-associated factors in the TFIID complex functionally interact with the TdT initiator element in a sequence-specific manner

The finding that TAFs are necessary for initiator-directed TATA-independent transcription from the TdT core promoter, independently of distal regulatory elements (see above), suggests a novel function of TAFs through specific sequences in the initiator element. To address this possibility more directly, we used an assay that takes advantage of the ability of the TdT initiator to synergize with a TATA element to markedly increase basal transcription (Smale *et al.*, 1990). This initiator function was confirmed by comparing the basal transcription activity in a HeLa nuclear extract of TdT promoter constructs (Figure 6A) having (i) only the initiator (TATA[−]/INR⁺), (ii) the initiator and a consensus TATA element (TATA⁺/INR⁺) or (iii) a consensus TATA element and a specifically mutated initiator (TATA⁺/

Fig. 6. TAFs are required for sequence-specific TdT initiator function in the absence of distal activators. (A) Structural features of core promoter constructs. The TATA[−]/INR⁺ promoter contains the TdT initiator element (boxed) and its natural flanking TATA-less TdT DNA sequences (thick line) from −41 to +33. The consensus AdML TATA element (Figure 1C) replaces natural TdT sequences from −25 to −31 in constructs TATA⁺/INR⁺ and TATA⁺/INR[−]. The underlined nucleotides of the mutant initiator in TATA⁺/INR[−] indicate the positions substituted. The arrow indicates the transcription start site. (B) Basal transcription activity in a HeLa nuclear extract (left panel, lanes 1–3) and in a TFIID-dependent transcription system reconstituted with purified factors and Pol II (right panel, lanes 4–11; see Materials and methods) of the constructs depicted in (A). In the right panel (lanes 4–11), + and − indicate promoters with the wild-type and the mutant initiator, respectively. No TFIID (lanes 4 and 5), 2 μ l native TFIID fraction (lanes 6 and 7) and 3 TBP U of either f:TBP (lanes 8 and 9) or f:TFIID (lanes 10 and 11) were used. An arrowhead indicates the correctly initiated transcripts.

INR[−]). As shown in Figure 6B, the insertion of a TATA box 30 bp upstream of the TdT initiator markedly increases basal transcription (lane 1 versus 2; see also Figure 1D). Most importantly, the mutation of specific base pairs important for initiator activity (Javahery *et al.*, 1994) in the TATA⁺/INR⁺ construct (to produce the TATA⁺/INR[−] construct), strongly decreases basal transcription activity (compare lane 2 with 3). These results confirm that TATA and initiator elements synergize in basal transcription and provide an assay to test for the requirement of TAFs in this initiator function; indeed, the maintenance of a low level of accurate initiation driven by the TATA element ensures that the initiator mutation does not simply destroy the ability of RNA polymerase to correctly initiate transcription. We then used a TFIID-dependent transcription system reconstituted with purified components (Figure 6B, right panel; see Materials and methods) to compare basal transcription activity of the TATA⁺/INR⁺ and TATA⁺/INR[−] constructs in the absence (lanes 4 and 5) and presence of either native partially purified TFIID (lanes 6 and 7), recombinant Flag epitope-tagged TBP (f:TBP; lanes 8 and 9) or highly purified Flag epitope-tagged TFIID (f:TFIID; lanes 10 and 11). As

observed with the crude nuclear extract, basal transcription driven by TFIID and the highly purified f:TFIID in this purified system is strongly decreased upon mutation (— lanes) of the initiator element (compare lanes 6 with 7 and 10 with 11). In contrast, transcription driven by f:TBP is not affected by the mutation in the initiator (lane 8 versus 9). Because of the use of an essentially homogeneous TFIID (f:TFIID) in the purified reconstituted system, this result thus provides the first evidence that TAFs can function through specific nucleotide positions in the initiator element to enhance distal activator-independent basal transcription.

Discussion

To address the mechanism of functional TBP recruitment to initiator-dependent TATA-less class II promoters, we have analyzed basal initiator-directed transcription from the natural TATA-less TdT promoter in the absence of functional upstream activator elements. Our results (Figure 1), together with previous analyses (Smale and Baltimore, 1989), demonstrate that efficient basal transcription from the natural TdT promoter in a HeLa nuclear extract does not depend on sequences upstream of the initiator, but requires only the well characterized initiator element (Smale *et al.*, 1990) and its flanking DR. Although the initiator is sufficient to direct low levels of specific transcription, the DR sequence contributes to efficient basal transcription activity in the absence of a TATA box. Interestingly, the DR sequence preferentially stimulates transcription in the absence of a TATA box (Figure 1D), thus distinguishing it from typical distal activator elements and from the recently described activator sequence that is present downstream of the HIV-1 transcription initiation site and which preferentially activates transcription *in vitro* in the presence of a TATA element (Zenzie-Gregory *et al.*, 1993b). Importantly, the TATA-less TdT initiator region can direct the formation of a stable (Sarkosyl- and heparin-resistant) preinitiation complex in the absence of a functional DR sequence (Figure 3). These results show clearly that the TATA-less TdT initiator region is sufficient for the functional recruitment of factors into a stable transcription preinitiation complex independently of distal promoter-bound activators. The DR sequence may facilitate the early initiator-directed TFIID recruitment step (see also below), although a possible function in other initiation or elongation steps cannot be excluded.

To address the mechanism of functional TFIID recruitment to the TATA-less TdT core promoter, we used a specific oligonucleotide competition-linked transcription assay. We demonstrated that blocking the TATA DNA-binding function of TFIID with a consensus TATA oligonucleotide inhibits transcription both from promoters containing strong (consensus) TATA elements and from a synthetic TdT initiator-dependent promoter with a weak non-consensus TATA-like sequence (pSp1 42 INR), but does not affect transcription from the natural TATA-less TdT and β -Pol promoters (Figure 2). This suggests that direct TBP binding to DNA contributes to efficient *in vitro* transcription from promoters with weak atypical TATA elements and explains the previously proposed functional requirement for TBP binding to the -30 DNA region of some natural and synthetic promoters with no consensus TATA sequence (Wiley *et al.*, 1992; Yoganathan *et al.*, 1992; Zenzie-Gregory *et al.*, 1993a). Most importantly, our analysis

provides the first evidence that the TATA DNA-binding activity of TBP is not required for the functional recruitment of TFIID to 'true' TATA-less promoters, like the natural TATA-less TdT and β -Pol promoters, and at the same time indicates that previous generalizations about TATA-less promoter functions based on the pSp1 42 INR template are potentially misleading. Moreover, we have shown that recombinant human TBP can promote specific basal transcription from TATA-containing promoters but not from the natural TATA-less TdT core promoter in a TFIID-dependent transcription assay (Figure 5). Similarly, we also observed that ectopic TBP stimulates TATA-dependent, but not TATA-independent, initiator-directed TdT transcription in a nuclear extract (data not shown). These results demonstrate that TBP is not limiting in the nuclear extract for basal TATA-less TdT initiator-directed Pol II transcription *in vitro* and correlate with the results of previous *in vivo* transfection experiments (Colgan and Manley, 1992). Altogether, these data further suggest that direct TBP–DNA interactions are not involved in the functional recruitment of TBP to the TATA-less TdT core promoter (Figure 7A).

Significantly, and in contrast to previous analyses using crude TFIID preparations and/or promoters with upstream activator DNA elements and cognate activating factors (Smale *et al.*, 1990; Pugh and Tjian, 1991; Zhou *et al.*, 1992), our highly purified native TFIID complex, composed of TBP and at least 13 specific tightly-associated polypeptides (TAFs; see Figure 4), can promote basal initiator-directed transcription in a TFIID-inactivated nuclear extract from the TATA-less TdT core promoter lacking distal activator sequences (Figure 5). Considering that neither TBP alone nor the TBP-containing TFIIB₂ complex can substitute for the highly purified TFIID complex used in this assay, and since the TATA-binding activity of TFIID is not required for its functional recruitment to the TdT promoter, this result strongly suggests a direct involvement of one or more TFIID-specific TAF(s) in the functional tethering of TBP to the TATA-less TdT initiator region in the absence of distal promoter-bound activators. Consistent with this view, our analysis of mutant TdT templates in a system reconstituted with purified factors has provided an important demonstration that TAFs are required for the basal (distal activator-independent) transcriptional stimulatory function of a wild-type initiator through specific initiator sequences (Figure 6).

Although we do not yet know the exact mechanism of TAF-dependent initiator-directed basal transcription, two general models can be considered (see Figure 7B). One or several TAFs could recruit TBP to a 'true' TATA-less promoter, either by interacting directly with the initiator and downstream DNA region (Figure 7B, middle) or by contributing to the tethering of TBP to an independent initiator-bound protein and/or Pol II (Figure 7B, right). In support of a possible direct interaction between TAFs and DNA over the initiator and downstream region is the previous observation that partially purified human and *Drosophila* TFIID generate extended DNase I footprints over the initiator and downstream regions of several promoters (Parker and Topol, 1984; Sawadogo and Roeder, 1985; Nakajima *et al.*, 1988; Emanuel and Gilmour, 1993; Purnell and Gilmour, 1993). Most importantly, a highly purified epitope-tagged TFIID complex can also generate the specific extended footprint covering the TATA box, the initiator and

Fig. 7. Models for TBP recruitment to a 'true' TATA-less promoter compatible with TAF- and initiator-dependent basal transcription. (A) TBP is tethered to a TATA-containing promoter by direct binding to the TATA box and promotes the assembly of a functional preinitiation complex (left). In contrast, the TATA-binding activity of TBP is not required and potential direct TBP–DNA interactions through domains distinct from the TATA-binding domain are not sufficient for the functional recruitment of TBP to a 'true' TATA-less promoter (right). (B) The TBP subunit of TFIID tethers the whole complex to a TATA-containing promoter (left). For initiator-dependent 'true' TATA-less core promoters, one or several TAFs are required for the functional recruitment of TBP to the core promoter region, either by direct interactions with the initiator and its downstream region (middle) or by interaction with an initiator-binding protein (IBP), which could be Pol II and/or a different initiator-binding factor (right). Considering the number of TAFs in TFIID and the large size of several of them (see Figure 4 and Chiang *et al.*, 1993), both the above mechanisms could contribute simultaneously to the recruitment of TBP to the TATA-less core promoter region (see text for discussion).

downstream sequences of the strong AdML promoter (Zhou *et al.*, 1992; Chiang *et al.*, 1993); in addition, this purified TFIID does not contain any of the previously identified initiator-binding proteins (i.e. HIP1, TFII-I, USF and YY1), as determined by molecular size and amino acid sequence comparisons with TAFs and by Western blot and DNA binding analyses (Chiang *et al.*, 1993; and data not shown). Furthermore, mutations in AdML initiator sequences that inhibit TFII-I and USF binding (Roy *et al.*, 1991; Du *et al.*, 1993) do not abolish protection over the initiator and downstream regions (Chiang *et al.*, 1993). In contrast, the epitope-tagged TFIID protects only the TATA box of the weaker HIV-1 and adenovirus E4 core promoters (Chiang *et al.*, 1993). Interestingly, the E4 promoter does not contain a TdT-like initiator sequence (Chiang *et al.*, 1993; and references therein), and in contrast to the TdT and AdML initiators, the HIV-1 initiator region is apparently not sufficient to direct basal transcription in the absence of a TATA box (Zenzie-Gregory *et al.*, 1993b). Taken together, these observations suggest a possible direct correlation between the function of the initiator region and its ability to direct the formation of extended TFIID footprints.

On the other hand, several lines of evidence also strongly suggest that the initial TFIID recruitment step may involve, or be enhanced by, a specific initiator-binding protein that is not present in the TFIID complex. For instance, both TFII-I and USF interact *in vitro* with the TdT initiator element

(Roy *et al.*, 1991). In addition, TFII-I also interacts with the TdT DR sequence (A.L.Roy and R.G.Roeder, unpublished results) which was shown here to be required for maximal basal transcription in the absence of a TATA box. Moreover, TFII-I can facilitate functional TBP recruitment to the AdML core promoter *in vitro*, most notably in the absence of TFIIA (Roy *et al.*, 1991, 1993a). Finally, it has been suggested that purified Pol II can also weakly recognize the initiator of AdML and AdIVa2 promoters (Carcamo *et al.*, 1991).

It should be emphasized, however, that the two models for initiator interactions are not mutually exclusive and that several simultaneous but weak interactions between different components in the TFIID complex and DNA, the other basal factors, an initiator-binding protein(s) and Pol II could probably stabilize the whole complex. This possibility is supported by the observation that the *in vitro* formation of a preinitiation complex on TATA-mutant AdML and AdIVa2 initiator regions is dependent on cooperative interactions between the basal transcription factors and Pol II (Carcamo *et al.*, 1991), by the demonstration of the cooperative assembly of alternate TFII-I-dependent preinitiation complexes on a TATA-mutated AdML and HIV-1 initiator regions (Roy *et al.*, 1993a) and by recent genetic and biochemical evidence for an interaction in yeast between the CTD domain of Pol II and a large TBP-containing multisubunit complex (Thomson *et al.*, 1993). In addition, although the TATA DNA-binding activity of TBP is not required for the initial recruitment of TFIID to 'true' TATA-less promoters (see above), it is formally possible that in the final preinitiation complex TBP may actually contact DNA in a sequence-independent manner and thus could also contribute to the overall stability of the complex. In this case, each of these potential particular interactions might contribute differentially in different core promoter contexts, as suggested, for instance, by our observation that the DNA-binding activity of TBP is required for efficient basal transcription from promoters with weak non-consensus TATA elements (see also Wiley *et al.*, 1992; Yoganathan *et al.*, 1992; Zenzie-Gregory *et al.*, 1993a) but not from 'true' TATA-less promoters, and by the proposed requirement for different initiator-binding proteins in different natural promoters (Weis and Reinberg, 1992, and references therein).

In contrast to previous studies on activator-dependent transcription using either the weak TATA-containing pSp1 42 INR construct or derivatives thereof (Smale *et al.*, 1990; Pugh and Tjian, 1991; Zhou *et al.*, 1992), our selective analysis of basal rather than activated transcription has allowed us to demonstrate a novel function of TAFs through the natural TATA-less initiator-dependent TdT core promoter. Indeed, our experiments with an essentially homogeneous TFIID demonstrate, for the first time, an upstream activator-independent function of TAFs through specific sequences in the initiator element and their involvement in mediating initiator-directed basal transcription from a 'true' TATA-less core promoter. Moreover, our TATA oligonucleotide competition experiments provide new evidence for important differences in the early steps of transcription initiation at TATA-dependent and initiator-dependent promoters (Zenzie-Gregory *et al.*, 1992) and, in addition, support the existence of alternative pathways for preinitiation complex assembly directed by TATA and

initiator elements (Roeder, 1991; Weis and Reinberg, 1992; Roy *et al.*, 1993a). Taken together, our results strongly suggest a possible alternative pathway for preinitiation complex assembly in which one or several TAFs contribute to TFIID recruitment to 'true' TATA-less class II core promoters by interacting with the initiator region either directly or indirectly through sequence-specific initiator binding proteins and/or Pol II (see above; Figure 7).

Interestingly, the largest of the TAFs (p230 in Figure 4) has been identified as a cell-cycle regulatory protein (CCG1) which contains a putative HMG box DNA-binding domain (Hisatake *et al.*, 1993; Ruppert *et al.*, 1993). This, plus the further demonstration that the *Drosophila* TAF homolog inhibits the TATA-binding activity of TBP *in vitro* (Kokubo *et al.*, 1993), suggest a potential contribution of TAF230/CCG1 in the regulation of TBP recruitment to TATA and/or initiator elements. In this respect, it is interesting to note that the anti-oncogene product p53 and the proto-oncoprotein Myc preferentially repress TATA-dependent and initiator/TFII-I-dependent transcription, respectively (Mack *et al.*, 1993; Roy *et al.*, 1993b). Finally, our results correlate with the recent observation of a requirement for Pol I- and Pol III-specific TAFs, interacting with core promoter elements either directly or indirectly through assembly factors, in recruiting TBP to TATA-less class I and class III genes, respectively (see Introduction), and thus underscore similarities in the mechanisms of TBP recruitment to TATA-less genes transcribed by all three nuclear RNA polymerases.

Materials and methods

Plasmid DNAs

The constructs pTdT(−1300/+59), pTdT(−1300/+33), pTdT(−41/+59) and pSp1 42 INR have been described previously (Smale and Baltimore, 1989) and were a kind gift from S. Smale. The plasmid pTdT(−41/+33) contains TdT promoter sequences from −41 to +33 obtained by PCR of pTdT(−41/+59) and cloned between the *EcoRI* and *BamHI* sites of pUC19. The PCR was carried out with *Pfu* polymerase (Stratagene) under conditions recommended by the manufacturer, with the 24mer M13/pUC primer number 1224 from New England Biolabs and the primer '+33/Bam': 3'-CTACCGTGTCTGTCTCTAGGGTACCAAGTC-5' having sequences complementary to TdT positions +19 to +33 (italics) and a *BamHI* site (bold). The PCR fragment cut with *EcoRI* and *BamHI* was inserted into pUC19 and sequenced. The plasmid pTdT(−41TATA/+59) was constructed by cloning between the *EcoRI* and *BamHI* sites of pUC19 an *EcoRI*–*BamHI* fragment obtained after double-digestion of the product of a PCR-mediated mutagenesis reaction on the template pTdT(−41/+59) with the mutagenic primer '−41/TATA': 5'-GCCAGTGAATTCGGTACCTATGTATAAA-ATGGTGAGAGGAC-3' having an *EcoRI* site (bold normal type) and a consensus TATA box (bold italic sequence) instead of the natural TdT sequences found between positions −25 and −31 (i.e. GGTCTGC) and the 24mer M13/pUC reverse primer number 1233 from New England Biolabs. The plasmid pTdT(−41TATA/+33) was obtained similarly by using the '−41/TATA' mutagenic primer and the '+33/Bam' primer for PCR. The constructs TATA⁺/INR⁺ and TATA⁺/INR[−] correspond to pTdT(−41/+33) and pTdT(−41TATA/+33), respectively. The construct TATA⁺/INR[−] was obtained by replacing the *EcoRI*–*BamHI* promoter fragment of TATA⁺/INR⁺ with a synthetic oligonucleotide having the same sequence but 6 bp substituted in the initiator element (see Figure 6A). All the constructs were confirmed by dideoxy sequencing. The plasmid pβP10 contains the human β-polymerase promoter (sequences −41 to +62) fused to the CAT gene (Widen *et al.*, 1988) and was obtained from S.H. Wilson. The construct pHsp70(−33/+99)–CAT was obtained from J.-B. Yoon and is a derivative of pHc(−33) (Simon *et al.*, 1988) by conversion of the *SphI* site to a *XhoI* site in the polylinker, and contains the human Hsp70 promoter (sequences −33 to +99) fused to the CAT gene. pE1b–CAT has been described previously (Lillie and Green, 1989). The plasmid pAdML(−257/+33)–CAT contains the adenovirus type 2 major late promoter sequences (from −257 to +33) fused to the CAT gene

and was provided by H. Kato. The bacterial expression vector for f:TBP, pF:TBP-11d, has been described previously (Chiang and Roeder, 1993; Chiang *et al.*, 1993). All plasmid DNAs were purified by two CsCl gradient centrifugations.

Nuclear extracts and proteins

HeLa nuclear extracts were prepared as described previously (Dignam *et al.*, 1983). Heat treatment of the nuclear extracts at 47°C was used to inactivate TFIID (Nakajima *et al.*, 1988). The length of the 47°C incubation period required to selectively inactivate TFIID without significantly affecting the other components of the general transcription machinery was found to vary for different nuclear extracts. The heat-treated nuclear extracts used in Figure 5 were inactivated for 6 min as described previously (Nakajima *et al.*, 1988). Pol II, TFIIA, TFIID, TFIIIE/F/H and USA were purified as described previously (Chiang *et al.*, 1993). Recombinant 6 histidine-tagged TFIIB was obtained from S. Malik (Malik *et al.*, 1991). Flag epitope-tagged human TBP (f:TBP; Figure 4, lane 3) was expressed in bacteria BL21(DE3)pLysS transformed with pF:TBP-11d and purified as reported (Chiang and Roeder, 1993). Flag epitope-tagged TBP–TAFs complexes f:TFIID and f:TFIIIB₂ were purified from nuclear extracts of the 3–10 HeLa cell line by phosphocellulose P11 (Whatman) chromatography followed by immunoaffinity on anti-Flag M2–agarose (IBI/Kodak) as described previously (Chiang *et al.*, 1993; see also the legend to Figure 4).

In vitro transcription and primer extension

Transcription reactions were carried out in a final volume of 20 µl and contained 40 or 80 fmol (see figure legends) of each supercoiled plasmid DNA template, 5 or 10 µl (see figure legends) of HeLa nuclear extract (of ~10 µg protein/µl), 20 mM HEPES (pH 7.9), 11–16% glycerol, 6 mM MgCl₂, 70 mM KCl, 5 mM DTT, 600 µM each of ATP, UTP, CTP and GTP, 40 U of recombinant RNasin (Promega) and 10 U of the RNA polymerase III inhibitor Tagetin (Epicentre Technologies). Reactions were incubated at 30°C for 60 min. For the oligonucleotide competition experiments in Figure 2, 5 µl of nuclear extract and 3.6 µl of double-stranded oligonucleotide competitor diluted in 10 mM Tris–HCl (pH 7.5), 100 mM NaCl and 1 mM EDTA (150-fold molar excess over the template, i.e. 12 pmol) were preincubated for 15 min at 30°C (in a final volume of 10 µl) before starting transcription by adding 10 µl of a mix containing all the other components including the template (80 fmol) and the nucleotides (see above). In the competition experiment of Figure 2B, the 10 µl mix containing the template was also preincubated in parallel for 15 min at 30°C in the presence of 2 µg BSA and with or without 10 ng of recombinant f:TBP before its addition to the oligonucleotide-blocked extract. The non-specific control oligonucleotide (NS) used has the sequence 5'-TCGAGCGGAGTA-CTGTCTCCGCTCGA-3' and the sequence of the oligonucleotide containing the AdML TATA box is 5'-CGATGGGGGGCTATAAAAGGG-GGTATCG-3' (only one strand of each double-stranded oligonucleotide is shown). For the transcription reactions with heat-inactivated nuclear extracts (Figure 5), the purified proteins and/or BSA (as indicated in the figure legend) were mixed with the extract containing all the other components (except DNA) on ice and transcription was started by the addition of the template and incubation at 30°C. Transcription reactions with purified factors contained 2 µl TFIIA, 30 ng 6hisTFIIB, 2 µl TFIID (or the indicated amount of f:TBP or f:TFIID; Figure 6B), 2 µl TFIIIE/F/H, 0.7 µl Pol II and 1 µl USA fraction in a 20 µl reaction containing 20 mM HEPES (pH 7.9), 70 mM KCl, 4 mM MgCl₂, 14% glycerol, 5 mM DTT, 0.5 µg/µl BSA, 600 µM of each rNTP and 40 U recombinant RNasin (Promega). After 60 min incubation at 30°C, all transcription reactions were stopped by adding 185 µl of stop solution containing 1% SDS, 5 mM EDTA, 150 mM NaCl, 20 mM Tris–HCl (pH 7.5), 20 µg glycogen and 40 µg Proteinase K, and incubation for 20 min at 37°C. RNA was extracted with phenol-chloroform, mixed with the kinased ³²P-labeled primer(s) (5' end-labeled; 1–2×10⁶ c.p.m., 0.5–1.0×10⁷ c.p.m./pmol) and precipitated with ethanol. We used the 16mer M13 Reverse Sequencing Primer number 1201 from New England Biolabs to prime RNAs produced by the TdT promoter-containing templates and the 30mer CAT primer 5'-GGTGGTATATCC-AGTGATTTTTTCTCCAT-3', which is complementary to the first 30 coding nucleotides of the CAT gene, to analyze the transcripts of promoter–CAT fusions. RNA and primer(s) were hybridized, after resuspension in 20 µl of hybridization buffer containing 10 mM Tris–HCl (pH 7.5), 250 mM KCl and 1 mM EDTA, for 90 min at either 55 (for CAT primer) or 37°C (for 16mer primer or when both primers were used simultaneously). The primer extension reaction was started by adding 40 µl of extension reaction solution containing 75 mM Tris–HCl (pH 8.0), 15 mM DTT, 12 mM MgCl₂, 75 µg/ml actinomycin D, 12 U recombinant RNasin (Promega), 750 µM each of dATP, dTTP, dCTP and dGTP, and 200 U of SuperScript RNaseH[−] reverse transcriptase (Gibco BRL), and

the reaction was incubated for 90 min at 37°C. The cDNA products were extracted with phenol-chloroform, precipitated and washed with ethanol, then dissolved and denatured (100°C for 4 min) in 6 µl of 95% formamide containing 20 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol FF, and finally analyzed on a 6% polyacrylamide–7.5 M urea sequencing gel.

Acknowledgements

We thank S.Smale, S.H.Wilson, J.-B.Yoon, H.Du and H.Kato for the gift of plasmids, S.Malik for the gift of 6hisTFIIB and J.Fondell, T.Oelgeschläger, K.Wang and J.-B.Yoon for critically reading the manuscript. This work was supported by National Institutes of Health grants CA42567 and AI27397 to R.G.R. and by general support from the Pew Trusts to the Rockefeller University. E.M. was supported by a postdoctoral fellowship of the Swiss National Science Foundation. C.-M.C. was a Helen Hay Whitney Foundation and Aaron Diamond Foundation postdoctoral fellow. H.G. was a Jane Coffin Childs Memorial Fund for Medical Research postdoctoral fellow.

References

- Bell, S.P., Jantzen, H.-M. and Tjian, R. (1990) *Genes Dev.*, **4**, 943–954.
- Bernués, J., Simmen, K.A., Lewis, J.D., Gunderson, S.I., Polycarpou-Schwarz, M., Moncollin, V., Egly, J.-M. and Mattaj, J.W. (1993) *EMBO J.*, **12**, 3573–3585.
- Breathnach, R. and Chambon, P. (1981) *Annu. Rev. Biochem.*, **50**, 349–383.
- Buratoski, S., Hahn, S., Sharp, P.A. and Guarente, L. (1988) *Nature*, **334**, 37–42.
- Carcamo, J., Buckbinder, L. and Reinberg, D. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 8052–8056.
- Chiang, C.-M. and Roeder, R.G. (1993) *Peptide Res.*, **6**, 62–64.
- Chiang, C.-M., Ge, H., Wang, Z., Hoffmann, A. and Roeder, R.G. (1993) *EMBO J.*, **12**, 2749–2762.
- Colgan, J. and Manley, J.L. (1992) *Genes Dev.*, **6**, 304–315.
- Comai, L., Tanese, N. and Tjian, R. (1992) *Cell*, **68**, 965–976.
- Cormack, B.P. and Struhl, K. (1992) *Cell*, **69**, 685–696.
- Cortes, P., Flores, O. and Reinberg, D. (1992) *Mol. Cell. Biol.*, **12**, 413–421.
- Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) *Nucleic Acids Res.*, **5**, 1475–1489.
- Du, H., Roy, A.L. and Roeder, R.G. (1993) *EMBO J.*, **12**, 501–511.
- Dynlacht, B.D., Hoey, T. and Tjian, R. (1991) *Cell*, **66**, 563–576.
- Emanuel, P.A. and Gilmour, D.S. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 8449–8453.
- Gregor, P.D., Sawadogo, M. and Roeder, R.G. (1990) *Genes Dev.*, **4**, 1730–1740.
- Hahn, S., Buratoski, S., Sharp, P.A. and Guarente, L. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 5718–5722.
- Hawley, D.K. and Roeder, R.G. (1985) *J. Biol. Chem.*, **260**, 8163–8172.
- Hawley, D.K. and Roeder, R.G. (1987) *J. Biol. Chem.*, **262**, 3452–3461.
- Hisatake, K., Hasegawa, S., Takada, R., Nakatani, Y., Horikoshi, M. and Roeder, R.G. (1993) *Nature*, **362**, 179–181.
- Hoffmann, A., Sinn, E., Yamamoto, T., Wang, J., Roy, A., Horikoshi, M. and Roeder, R.G. (1990) *Nature*, **346**, 387–390.
- Huet, J. and Sentenac, A. (1992) *Nucleic Acids Res.*, **20**, 6451–6454.
- Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B. and Smale, S.T. (1994) *Mol. Cell. Biol.*, **14**, 116–127.
- Jiang, Y., Smale, S.T. and Gralla, J.D. (1993) *J. Biol. Chem.*, **268**, 6535–6540.
- Kassavetis, G.A., Joazeiro, C.A.P., Pisano, M., Geiduschek, E.P., Colbert, T., Hahn, S. and Blanco, J.A. (1992) *Cell*, **71**, 1055–1064.
- Kokubo, T., Gong, D.-W., Yamashita, S., Horikoshi, M., Roeder, R.G. and Nakatani, Y. (1993) *Genes Dev.*, **7**, 1033–1046.
- Learned, R.M., Learned, T.K., Haltinger, M.M. and Tjian, R. (1986) *Cell*, **45**, 847–857.
- Lillie, J.W. and Green, M.R. (1989) *Nature*, **338**, 39–44.
- Lo, K., Landau, N.R. and Smale, S.T. (1991) *Mol. Cell. Biol.*, **11**, 5229–5243.
- Lobo, S.M., Lister, J., Sullivan, M.L. and Hernandez, N. (1991) *Genes Dev.*, **5**, 1477–1489.
- Lobo, S.M., Tanaka, M., Sullivan, M.L. and Hernandez, N. (1992) *Cell*, **71**, 1029–1040.
- Mack, D.H., Vartikar, J., Pipas, J.M. and Laimins, L.A. (1993) *Nature*, **363**, 281–283.
- Malik, S., Hisatake, K., Sumimoto, H., Horikoshi, M. and Roeder, R.G. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 9553–9557.
- Margottin, F., Dujardin, G., Gerard, M., Egly, J.-M., Huet, J. and Sentenac, A. (1991) *Science*, **251**, 424–426.
- Means, A.L. and Farnham, P.J. (1990) *Mol. Cell. Biol.*, **10**, 653–661.
- Means, A.L., Slansky, J.E., McMahon, S.L., Knuth, M.W. and Farnham, P.J. (1992) *Mol. Cell. Biol.*, **12**, 1054–1063.
- Meisterernst, M., Roy, A.L., Lieu, H.M. and Roeder, R.G. (1991) *Cell*, **66**, 981–993.
- Nakajima, N., Horikoshi, M. and Roeder, R.G. (1988) *Mol. Cell. Biol.*, **8**, 4028–4040.
- O'Shea-Greenfield, A. and Smale, S.T. (1992) *J. Biol. Chem.*, **267**, 1391–1402.
- Parker, C.S. and Topol, J. (1984) *Cell*, **36**, 357–369.
- Parvin, J.D. and Sharp, P.A. (1993) *Cell*, **73**, 533–540.
- Peterson, M.G., Tanese, N., Pugh, B.F. and Tjian, R. (1990) *Science*, **248**, 1625–1630.
- Pugh, B.F. and Tjian, R. (1990) *Cell*, **61**, 1187–1197.
- Pugh, B.F. and Tjian, R. (1991) *Genes Dev.*, **5**, 1935–1945.
- Purnell, B.A. and Gilmour, D.S. (1993) *Mol. Cell. Biol.*, **13**, 2593–2603.
- Reinberg, D. and Roeder, R.G. (1987) *J. Biol. Chem.*, **262**, 3310–3321.
- Reinberg, D., Horikoshi, M. and Roeder, R.G. (1987) *J. Biol. Chem.*, **262**, 3322–3330.
- Roeder, R.G. (1991) *Trends Biochem. Sci.*, **16**, 402–408.
- Roy, A.L., Meisterernst, M., Pognonec, P. and Roeder, R.G. (1991) *Nature*, **354**, 245–248.
- Roy, A.L., Malik, S., Meisterernst, M. and Roeder, R.G. (1993a) *Nature*, **365**, 355–359.
- Roy, A.L., Carruthers, C., Gutjahr, T. and Roeder, R.G. (1993b) *Nature*, **365**, 359–361.
- Ruppert, S., Wang, E.H. and Tjian, R. (1993) *Nature*, **362**, 175–179.
- Sawadogo, M. and Roeder, R.G. (1985) *Cell*, **43**, 165–175.
- Schultz, M.C., Reeder, R.H. and Hahn, S. (1992) *Cell*, **69**, 697–702.
- Seto, E., Shi, Y. and Shenk, T. (1991) *Nature*, **354**, 241–245.
- Shi, Y., Seto, E., Chang, L.-S. and Shenk, T. (1991) *Cell*, **67**, 377–388.
- Simmen, K.A., Bernues, J., Parry, H.D., Stunnenberg, H.G., Berkenstam, A., Cavallini, B., Egly, J.-M. and Mattaj, J.W. (1991) *EMBO J.*, **10**, 1853–1862.
- Simmen, K.A., Bernues, J., Lewis, J.D. and Mattaj, J.W. (1992) *Nucleic Acids Res.*, **20**, 5889–5898.
- Simon, M.C., Fisch, T.M., Benecke, B.J., Nevins, J.R. and Heintz, N. (1988) *Cell*, **52**, 723–729.
- Singer, V., Wobbe, R. and Struhl, K. (1990) *Genes Dev.*, **4**, 636–645.
- Smale, S.T. and Baltimore, D. (1989) *Cell*, **57**, 103–113.
- Smale, S.T., Schmidt, M.C., Berk, A.J. and Baltimore, D. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 4509–4513.
- Taggart, A.K.P., Fisher, T.S. and Pugh, B.F. (1992) *Cell*, **71**, 1015–1028.
- Takada, R., Nakatani, Y., Hoffmann, A., Kokubo, T., Hasegawa, S., Roeder, R.G. and Horikoshi, M. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 11809–11813.
- Tanese, N., Pugh, B.F. and Tjian, R. (1991) *Genes Dev.*, **5**, 2212–2224.
- Thomson, C.M., Koleske, A.J., Chao, D.M. and Young, R.A. (1993) *Cell*, **73**, 1361–1375.
- Weis, L. and Reinberg, D. (1992) *FASEB J.*, **6**, 3300–3309.
- White, R.J. and Jackson, S.P. (1992) *Cell*, **71**, 1041–1053.
- White, R.J., Jackson, S.P. and Rigby, P.W.J. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 1949–1953.
- Widen, S.G., Kedar, P. and Wilson, S.H. (1988) *J. Biol. Chem.*, **263**, 16992–16998.
- Wiley, S.R., Kraus, R.J. and Mertz, J.E. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 5814–5818.
- Yoganathan, T., Horikoshi, M., Hasegawa, S., Roeder, R.G. and Sells, B.H. (1992) *Biochem. J.*, **285**, 721–723.
- Zawel, L. and Reinberg, D. (1992) *Curr. Opin. Cell Biol.*, **4**, 488–495.
- Zenzie-Gregory, B., O'Shea-Greenfield, A. and Smale, S.T. (1992) *J. Biol. Chem.*, **267**, 2823–2830.
- Zenzie-Gregory, B., Khachi, A., Garraway, I.P. and Smale, S.T. (1993a) *Mol. Cell. Biol.*, **13**, 3841–3849.
- Zenzie-Gregory, B., Sheridan, P., Jones, K.A. and Smale, S.T. (1993b) *J. Biol. Chem.*, **268**, 15823–15832.
- Zhou, Q., Lieberman, P.M., Boyer, T.G. and Berk, A.J. (1992) *Genes Dev.*, **6**, 1964–1974.

Received on March 10, 1994; revised on April 11, 1994