Non-optimal TATA Elements Exhibit Diverse Mechanistic Consequences*

Jennifer J. Stewart, Julie A. Fischbeck, Xu Chen, and Laurie A. Stargell

From the Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523-1870

To reveal mechanistic differences in transcription initiation between variant TATA elements, in vivo and in vitro assays of the functional activity of 14 different sequences were compared. Variant elements exhibited particular degrees of activation in vivo but universally were unable to support the -fold activation observed for an element consisting of TATAAA. Each element classified by its functional activity for in vitro interaction with TATA-binding protein (TBP), TFIIA, and TFIIIB. Certain off-consensus TATA elements form poor binding sites for TBP and this compromised interaction interferes with higher order complex formation with TFIIA and/or TFIIIB. Other elements are only modestly decreased for TBP binding but dramatically affected for higher order complex formation. Another distinct category is comprised of two elements (CATAAA and TATAAG), which are not affected in the initial formation of the TBP, TFIIA-TBP, or TFIIIB-TBP complexes. However, CATAAA and TATAAG are unable to form a stable TFIIA-TBP-DNA complex in vitro. Moreover, fusion of TFIIA to TBP specifically restores activity from these two elements in vivo. Taken together, these results indicate that the interplay between the sequence of the TATA element and the components of the general transcription machinery can lead to variations in the formation of functional complexes and/or the stability of these complexes. These differences offer distinct opportunities for an organism to exploit diverse steps in the regulation of gene expression depending on the precise TATA element sequence at a given gene.

Initiation of transcription by RNA polymerase II is the major site for regulation of eukaryotic gene expression during cell cycle progression, development, and physiological induction (1). Increased gene expression is mediated by activator proteins that bind regulatory sequences in the promoter and, either directly or indirectly, facilitate recruitment of the general transcription factors assembles near the initiation site at the core promoter (2). TATA-binding protein (TBP) specifically recognizes and binds to the TATA sequence of the core promoter allowing for the nucleation of other general transcription factors including TFIIA, -B, -E, -F, -H, -J, and polymerase II (for reviews, see Refs. 3 and 4). Although high affinity binding sites have been identified for yeast TBP (5–7), and computational studies indicate that TATA(A/T)(A/T)(A/G) is commonly found at TBP-dependent promoters in yeast (8), a wide variety of off-consensus elements have been shown to exhibit transcriptional activity. Indeed, these alterations from the consensus sequence are not passive players in gene regulation but can contribute to the proper expression of a given gene. For example, activation by E1A of the human hsp70 promoter is highly dependent on the sequence of the TATA element: substitution of the hsp70 TATA (TATAA) with the SV40 early promoter TATA (TATTTAT) results in loss of induction by E1A but not by heat shock (9). In addition, an off-consensus TATA element is essential for the proper regulation of the osteocalcin gene by glucocorticoid receptor (10). As such, the particular sequence of the TATA element is likely to be critical for appropriate gene expression of a specific gene.

To understand how variations in the TATA element affect transcription initiation, we compared the functional activity of 14 different elements. Because the first six bases are the most well defined in the 8-base pair sequence, we focused our efforts on replacements within these bases (positions 1 through 6). Each position of the first six bases was replaced with either a cytosine (C) or guanine (G) base. These elements were tested for their ability to direct transcription in vivo, and to form complexes with TBP, TFIIA, and TFIIIB in vitro. The use of identical sequence context, recombinant proteins, and conditions, allows for a detailed dissection of the mechanistic role that TATA sequence variation plays in transcription initiation. We found that each of the substituted elements is reduced for directing high levels of gene expression in vivo compared with TATA. Electrophoretic mobility shift assays with TBP, TFIIA, and TFIIIB demonstrate that many of these TATA variants are compromised for formation of particular complexes, indicating sequence-specific alterations in complex formation. Two elements, TATAAG and CATAAA, fell into a previously described category of TATA variants (11) that are compromised for the stability of the TFIIA-TBP-DNA complex, whereas other complexes are unaffected. Overall, the data suggest that different positions of the TATA box can influence the kinetics of factor binding and/or the stability of the complex, which could contribute to multiple opportunities for transcription regulation. This work represents the first description of the functional activity of variant TATA elements for higher order complex formation and serves as the foundation for further...
TATA Element Variant Analyses

studies aimed at understanding the mechanistic penalties associated with non-consensus TATA elements.

EXPERIMENTAL PROCEDURES

Transcriptional Analyses—Plasmids used for measuring activated transcription in vivo were derivatives of Ycp86 containing the hybrid HIS3 promoter and the wild-type initiation and amino-terminal region of the HIS3 gene fused in-frame with functional Escherichia coli LacZ (5, 6). The promoter region contains a 365-bp GAL1,10 fragment containing four GAL4 binding sites fused upstream of the EcoRI-SacI restriction endonuclease sites, between which the substituted element oligonucleotides could be inserted. When the synthesized oligos were cloned into this molecule, the plasmids were renamed pJS3801 through pJS3814 (Table 1).

For the experiments measuring gene expression in vivo, the pJS3801–pJS3814 constructs were transformed into yeast strain 5YS156, a derivative of KY804 (relevant genotype: MATα ura3–52 trp1–Δ1 leu2–Δ2::PET56 gal2 gcv4–Δ1) (12) with the chromosomal copy of SPT15 deleted by a two-step knock-out. TBP functions are provided by a URA3-marked plasmid containing the TBP promoter, open reading frame, and terminator. Plasmid shuffling on 5-fluoroorotic acid was used to create the hybrid TBP promoter, open reading frame, and terminator. Plasmids used for measuring transcription in vivo were used in liquid assays for functional amino-terminal region of the hybrid TBP, and the resulting transformants were equilibrated in 100 mM KCl, 40 mM HEPES, 20 mM Tris, 10% glycerol, 1 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride (final pH 7.9). The protein expression and purification. The TFIIA was eluted using 5 mM reduced glutathione in 50 mM Tris borate, 1 mM EDTA, and 2 mM MgCl2 in 100 mM KCl, 40 mM KCl, 20 mM Tris, 10% glycerol, 1 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride (final pH 7.9). The protein was shown to be 90% pure upon staining with Coomassie Blue.

Protein Expression and Purification—Full-length yeast TBP was expressed in E. coli strain BL21 DE3 (13). The protein was purified from the soluble fraction with Q, SP, and Heparin HiTrap columns (Amersham Biosciences). The resulting fractions were equilibrated in 100 mM KCl, 40 mM HEPES, 20 mM Tris, 10% glycerol, 1 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride (final pH 7.9). The protein was shown to be 90% pure upon staining with Coomassie Blue.

TFII B was produced by fusing the open reading frame to glutathione S-transferase and transformed into E. coli BL21 DE3 (Novagen). Cells were grown in LB medium at 37 °C to an optical density of 0.7. Isopropyl β-D-thiogalactopyranoside (0.1 mM final concentration) was added and the cells were incubated for 2 h at 30 °C. Cells were harvested by centrifugation and washed with a 20 mM Tris, 50 mM NaCl buffer. Following sonication, the lysate was incubated with shaking at 4 °C with glutathione resin. After two consecutive wash steps, the protein was eluted using 5 mM reduced glutathione in 50 mM Tris buffer. The eluate was equilibrated in 100 mM KCl, 40 mM HEPES, 20 mM Tris, 10% glycerol, 1 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride (final pH 7.9). The protein was shown to be 75% homogeneous on SDS-PAGE upon staining with Coomassie Blue.

Recombinant yeast TFIIIA was purified as described (14). This procedure involves expressing each subunit, TOA1 and TOA2, in separate strains of E. coli BL21 DE3. Cells were ruptured by sonication; insoluble material was collected by centrifugation. Each insoluble pellet was resolubilized in 8 M urea.

Each subunit was renatured in the presence of the other subunit and dialyzed against the buffer described in the TBP and TFII B purification. The TFIIA was ~60% pure as determined by Coomassie staining.

Electrophoretic Mobility Shift Assays (EMSA)—The DNA elements used for the in vitro protein DNA interaction studies were 23-base pair oligonucleotides (11). The TATAAA oligo is 5′-AATTCCCTATAAAAGTAAATGAGGAG-3′. The other elements were synthesized with the appropriate base substitution and prepared in the exact same way.

Protein-DNA interactions in vitro were studied by incubation of purified proteins with 32P internally labeled probes. Binding reactions contained 10 μM poly(dG-dC) nonspecific competitor, 100 mM KCl, 40 mM HEPES, 20 mM Tris, 10% glycerol, 1 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride (final pH 7.9). TBP (13.8 nM) and the TATA element probe (2.4 pmol) were incubated for 30 min at room temperature and the complex was separated on a 5% acrylamide gel containing 50 mM Tris borate, 1 mM EDTA, and 2 mM MgCl2, in both the gel and running buffer. Recombinant yeast TFIIA (5.0 nM) and TFII B (8.5 nM) were incubated with TBP and probe DNA as described above except that MgCl2 was omitted from the gel and running buffer. Omission of the divalent cations from the binding reaction allows for the stable formation of

### TABLE 1

DNA oligonucleotides used in this study

| Name         | Sequence*     |
|--------------|---------------|
| pJS801       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS803       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS804       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS805       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS806       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS807       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS808       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS809       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS810       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS811       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS812       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS813       | GAATTCCATAAAAATTAAGTGAGCTC |
| pJS814       | GAATTCCATAAAAATTAAGTGAGCTC |

* The sequence of the synthesized oligonucleotides (23 bases) are shown. The six bases that were the focal point of the study are indicated in bold, and the underlines highlight the replacement within each particular derivative.

### TABLE 2

TATA variants are reduced for activated transcription in vivo

| Element* | Raffinose -Fold Activation | Galactose -Fold Activation |
|----------|---------------------------|----------------------------|
| TATAAA   | 1.8                       | 100                        |
| CATAAA   | 1.4                       | 5.4                        |
| TACAAA   | 2.1                       | 4.4                        |
| TATCAA   | 1.5                       | 4.7                        |
| TATACA   | 1.8                       | 4.8                        |
| TATAC   | 1.5                       | 2.8                        |
| GATAAA   | 1.1                       | 3.2                        |
| TGTAAA   | 1.7                       | 3.3                        |
| TAGTAAA  | 1.9                       | 3.2                        |
| TAGTAA   | 1.6                       | 3.2                        |
| TATAAG   | 1.5                       | 3.7                        |

* β-Galactosidase activities were performed with cells cultured in medium containing either 2% raffinose or 2% galactose for 18 h. Values were set to TATAAA in galactose equaling 100%. Data were averages from three independent experiments and standard errors were within 4–6% of a given value.

### Notes

- β-Galactosidase activities were performed with cells cultured in medium containing either 2% raffinose or 2% galactose for 18 h. Values were set to TATAAA in galactose equaling 100%. Data were averages from three independent experiments and standard errors were within 4–6% of a given value.
- The yeast strain yJS156 was transformed with a HIS3/Lac Z fusion plasmid containing either the canonical TATAAA element or the indicated non-canonical element, driven by the GAL1,10 UAS.

Each subunit was renatured in the presence of the other subunit and dialyzed against the buffer described in the TBP and TFII B purification. The TFIIA was ~60% pure as determined by Coomassie staining.
only the ternary complex. For dissociation kinetic studies, TBP-DNA, TFIIA-TBP-DNA, and TFIIB-TBP-DNA complexes were incubated for 30 min and then challenged with 1000-fold molar excess of specific competitor poly(dA-dT) for the specified amount of time. The samples were then loaded on to the gel and resolved by EMSA.

RESULTS

Single Cytosine or Guanine Replacements in the TATAAA Element Diminish Gene Expression in Vivo—A set of variant elements was designed based on the first six bases of the natural HIS3 TATA element (TATAAAGT; Table 1). Variants of the TATA element with cytosine bases substituted at each of the first six positions were termed the C series. Because there was a concern about the effect of substituting a purine (adenine) for a pyrimidine (cytosine), a series of guanine base substitutions at each of the first six positions was also constructed (the G series). For all of the variants, the flanking sequences remained identical. To test for transcriptional capability in vivo, each element was cloned individually into the core promoter region of a reporter plasmid containing the GAL1, 10 UAS and the HIS3 initiation region driving the expression of a HIS3-LacZ fusion (5, 6). This reporter system was used to measure the response to Gal4, a potent acidic activator that stimulates transcription in the presence of galactose. Incubation in galactose medium resulted in a greater than 50-fold induction in \( \Delta \beta \)-galactosidase activity for cells containing the TATAAA driven reporter (Table 2). In contrast, the reporters driven by the non-canonical elements exhibited inductions ranging from less than 2-fold to nearly 4-fold, depending on the particular sequence. To deter-

FIGURE 1. The formation of the TBP-DNA complex on the C and G series of non-canonical elements is sequence dependent. Representative EMSA experiments showing recombinant yeast TBP incubated with each of the 14 \( 32 \)P-labeled probes in the C series (panel A) and the G series (panel B). Free DNA is denoted by an asterisk and the TBP-DNA complex is indicated by arrows.

FIGURE 2. The formation of higher order complexes on the C and G series of non-canonical elements shows sequence-dependent binding defects. Representative EMSA experiments showing recombinant yeast TBP, TFIIA, and TFIIB incubated with the \( 32 \)P-labeled probes with the C series (panel A) and G series (panel B) variants. Free DNA is denoted by an asterisk, whereas the TFIIA-TBP-DNA and TFIIB-TBP-DNA complexes are indicated by the appropriate arrows.

FIGURE 3. Complex formation on C- and G-substituted non-canonical elements is sequence dependent. Experimental results from Figs. 1 and 2 are presented graphically to summarize the degree of protein complex formation on the C series (panel A) and G series (panel B) of non-canonical elements. The total amount of binding to the canonical TATA element is shown as a function of the position of the substitution in each element. Data are mean ± the range from two independent experiments.
mine the impact of altering the TATA sequence on protein-DNA interactions, the extent of TBP-DNA complexes formed on each element was determined in vitro.

Internal Base Substitutions in the TATA Sequence Can Disrupt TBP Binding—Each of the C series elements was tested for the ability to form the TBP-DNA complex (Fig. 1A). As reported in previous work (11), substitution of the first T position with a C (CATAAA) did not alter the amount of TBP-DNA complex formed compared with TATAAA. Two other elements (TACAAA and TATAAC) also exhibited fairly robust TBP binding (60–70% of TATAAA activity). Substitutions at the second, fourth, and fifth positions with C (TCTAAA, TATCAA, and TATACA) were not as well tolerated, with less than 50% of the TBP-DNA complex formed on TATAAA or CATAAA.

TBP-DNA complex formation was also measured on the G series elements (Fig. 1B). Interestingly, a substitution of the sixth position A with G (TATAAG) formed the TBP-DNA complex to the same degree as TATAAA and CATAAA, indicating that retaining a purine at this position is better tolerated than a pyrmidine (compare with TATAAC in Fig. 1). This was also observed for the first and third positions: compare the activity of CATAAAA (nearly 100%) to GATAAA (50%); and TACAAAA (65%) to TATGAA (45%). The TATGAA and GATAAA substitutions were the most functional, showing 50% of TATAAA and CATAAA binding. In contrast, the TATAGA element showed the most compromised ability to form the TBP-DNA complex, with less than 30% of the amount of complex formed on the control elements.

Higher Order Complex Formation Also Depends on the Sequence of the Core Element—Higher order complex formation was measured for the C series of elements and compared with the binding to the TATAAA element (Fig. 2A). TFIIA-TBP-DNA complex formation on the elements ranged from TCTAAA with 53% of TATAAA binding, to TATCAA showing the most dramatic decrease in TFIIA-TBP-DNA binding (25%). A similar trend was observed with the formation of the TFIIB-TBP-DNA complex on the C series of elements, although TFIIB complexes seem to be more significantly affected by the cytosine base substitutions with activities ranging from 12 to 35% of TATAAA binding (Fig. 2A and 3A).

In the case of TFIIA-TBP-DNA complex formation, the TATAAG element formed this complex to a similar extent as TATAAA. The TATGAA element is an interesting case because this element is compromised for both TBP and TFIIA-TBP complex formation but not TFIIB-TBP-DNA complex formation, suggesting a TFIIB-dependent conformational change is stabilizing the TBP-DNA interaction. Conversely, the GATAAA element is much more functional in forming the TBP-DNA and TFIIA-TBP-DNA complexes.
(50–65% TATAAA binding), but is extremely defective for formation of the TFIIA-TBP-DNA complex (5%). The remaining elements do not show significant TFIIA-TBP-DNA complex formation, with amounts of complex formed at \( \leq 10\% \) of TATAAA (Figs. 2B and 3B).

**The TFIIA-TBP-DNA Complex Is Destabilized on the TATAAG Element**—In the studies presented above, many of the elements have specific alterations in protein-DNA complex formation. However, certain substitutions of the first (with a C) and sixth (with a G) positions do not appear to affect initial complex formation of TBP-DNA, TFIIA-TBP-DNA, or TFIIIB-TBP-DNA complexes. We have demonstrated previously that the CATAAA element is unable to form a stable TFIIA-TBP-DNA complex (11). To test the hypothesis that TATAAG may have a similar affect on TFIIA-TBP-DNA complex stability, electrophoretic mobility shift assays were used to measure the relative binding and stability of the TFIIA-TBP complex on both TATAAA and TATAAG (Fig. 4A). A comparison of the absolute amount of binding of TFIIA-TBP to each of the elements shows very similar amounts of complex formation on both elements. The stability of the TFIIA-TBP-DNA complex was determined by examining the dissociation kinetics for the two complexes. The stability of the TFIIA-TBP-TATA complex on TATAAA differs significantly from the complex formed on the TATAAG element. The TFIIA-TBP complex was extremely stable on TATAAA, with little or no loss of complex over the 2-h time course. In contrast, the TFIIA-TBP complex on TATAAG showed significant loss of complex, with almost complete loss of complex over the 2-h time course.

Because the stability of the TFIIA-TBP-DNA complex is compromised on the TATAAG element, we wished to determine whether this was specific to the TFIIA-TBP complex, or whether a loss of stability is also observed for other complexes. The TBP-DNA complexes form and decay with a similar kinetic rate on both TATAAA and TATAAG (Fig. 4B). The TFIIA-TBP-DNA complex also behaved very similarly on both TATAAA and TATAAG as well (Fig. 4C). In conclusion, TBP binds to TATAAA and TATAAG in such a manner that allows for the stable association of the TFIIA-TBP higher order complex and the difference in stability of the TFIIA-TBP-DNA complex is dependent on the sequence of the TATAAA element.

**FIGURE 5.** The stability of the TFIIA-TBP-DNA complex is dependent on the sequence of the TATAAA element. Representative EMSAs showing the binding and stability of the TFIIA-TBP complex on different classes of core promoter elements: elements that, like TATAAA, exhibit infinite stability (panel A); elements that exhibit slight loss of complex over time (panel B); and elements that show formation of very unstable TFIIA-TBP-DNA complexes (panel C). In each case, TBP and TFIIA were incubated with the DNA elements for 30 min before the dissociation time courses were initiated by adding specific competitor (lane 3). Each subsequent lane represents an additional 30 min of incubation with specific competitor. Controls of DNA probe alone (lane 1) and TBP without the addition of TFIIA in the absence of Mg\(^{2+}\) (lane 2) are also shown. Free DNA is indicated by an asterisk and the TFIIA-TBP-DNA complex is indicated by a solid arrow. At right, graphical representation of the exponential decay of the TFIIA-TBP-DNA complexes (symbols correspond to those shown next to the DNA sequence under each EMSA). Complex remaining at each time point is plotted as fraction remaining versus time. Data are mean \( \pm \) the range of two independent experiments.
complex on TATAAG, like CATATAA (11), is specific to the complex containing TFIIA.

Lack of Stability of the TFIIA-TBP-DNA Complex Is Also Observed for Other Elements—The stability of the TFIIA-TBP-DNA complex on the remaining TATAAA variants was also tested. Although the TATCAA, TGTAAG, and TAGAAA elements do not form TFIIA-TBP-DNA complex to the same extent as TATAAA, the amount of complex that does form is stable over the course of the experiment (Fig. 5 A). Each subsequent lane represents an additional 30 min of incubation with specific competitor. Controls of DNA probe alone (lane 1) and TBP without the addition of TFIIA in the absence of Mg$^{2+}$ (lane 2) are also shown. Free DNA is indicated by an asterisk and the TFIIA-TBP-DNA complex is indicated by a solid arrow. At right, graphical representation of the exponential decay of the TFIIA-TBP-DNA complexes (symbols correspond to those shown next to the DNA sequence under each EMSA). Complex remaining at each time point is plotted as fraction remaining versus time. Data are mean ± the range of two independent experiments.

The Stability of the TFIIA-TBP-DNA Complex Depends Strongly on the Sequence of the TATAAA element—Competition studies were conducted to measure the stability of the TFIIA-TBP-DNA on the TATAAA variants. Although the TATAAC, TACAAA, and TCTAAA elements do not form the TFIIA-TBP-DNA complex to the same extent as TATAAA, the kinetics of the loss of complex from these elements is very similar to that of TATAAA (Fig. 6 A). The TGTAAA, TAGAAA, and TATGAA elements form initial amounts of the TFIIA-TBP-DNA complex, but these elements share the trend that there is nearly complete loss of the complex very quickly during the competition, usually in less than 30 min (Fig. 6 B). The final group containing elements TATCAA, TATACA, and GATAAA form very low initial amounts of the TFIIA-TBP-DNA complex, and these complexes are as stable as the TFIIA-TBP-TATAAA complex (Fig. 6 C).

Fusion of TBP and TFIIA Results in an Increase in Expression from CATATAAA and TATAAG in Vivo—The in vitro results suggest that the lack of formation of stable protein-DNA complexes could contribute to the extremely low levels of gene expression from these elements. The stability of the TFIIA-TBP-DNA complex is influenced by the sequence of the TATAAA element. Representative EMSAs showing the binding and stability of the TFIIA-TBP complex on different classes of core promoter elements: elements that are similar to TATAAA in their dissociation kinetics (panel A); elements that exhibit extremely rapid loss of complex over time (panel B); and elements that form less initial TFIIA-TBP-DNA complex, but that which forms appears to be stable (panel C). In each case, TBP and TFIIA were incubated with the DNA elements for 30 min before the dissociation time courses were initiated by adding specific competitor (lane 3). Each subsequent lane represents an additional 30 min of incubation with specific competitor. Controls of DNA probe alone (lane 1) and TBP without the addition of TFIIA in the absence of Mg$^{2+}$ (lane 2) are also shown. Free DNA is indicated by an asterisk and the TFIIA-TBP-DNA complex is indicated by a solid arrow. At right, graphical representation of the exponential decay of the TFIIA-TBP-DNA complexes (symbols correspond to those shown next to the DNA sequence under each EMSA). Complex remaining at each time point is plotted as fraction remaining versus time. Data are mean ± the range of two independent experiments.
expression observed from several of the non-canonical elements in vivo. One would predict that if these complexes could be stabilized in vivo, this would result in an increase in transcriptional activity. To test this hypothesis, we utilized yeast strains expressing TFIIA-TBP and TFIIB-TBP fusion molecules. Presumably, fusing TBP directly to a subunit of TFIIA (Toa2) or to TFIIIB will artificially increase the effective concentration of these proteins at the promoter element being tested.

The response to Gal4 in the strain containing the TFIIA-TBP or TFIIB-TBP fusion as the sole source of TBP was measured for all 13 elements using the β-galactosidase assay. Incubation in galactose medium for 18 h resulted in significant induction in activity in both the TFIIA-TBP and TFIIB-TBP fusion strains harboring the TATAAA driven reporter (Fig. 7), as expected for a canonical element. Strikingly, the TFIIA-TBP fusion strains containing the CATATAA and TATAAG driven reporters showed a dramatic increase in activity to ~30% of TATAAA activity. This increase in CATATAA and TATAAG activity is specific to these elements because none of the other elements showed any increase in activity in the Toa2-TBP fusion strains. Moreover, the increase is specific for TFIIA because no such increase is observed in the TFIIB-TBP fusion strain. Thus, transcriptional output from elements characterized by an unstable TFIIA-TBP-DNA complex can be specifically enhanced by fusing TFIIA and TBP.

**DISCUSSION**

Recent studies using whole genome analysis have shifted the view of a rigid requirement for the TATA element for transcriptional activity from genes expressed by RNA polymerase II. In fact, it is estimated that less than one-third of promoters actually contain sequences that resemble the classical TATA motif (8, 15–17). Inherent flexibility in the TBP-DNA interaction is also suggested by the observations that yeast TBP binds similarly to TATAAA and CATATAA elements (11), and Arabidopsis TBP binds 10 variations of the adenovirus major late promoter TATA element in a structurally similar manner (18). And yet, even subtle sequence variations that still fall within the consensus sequence can alter TBP-TATA interactions (19). However, the observations that the absolute amount of TBP binding to off-consensus elements does not always correlate well with transcriptional activity (20–22), and that consensus strong elements can possess markedly distinct stepwise interactions with TBP (23), indicate that there may be significant consequences of alternative TBP-DNA complexes. Taken together, these studies suggest that the TBP-DNA complex may exist in different structural contexts for variant sequences, and this context may also impact downstream steps in the transcription process. Because variant TATA elements play a fundamental role in the proper expression of many eukaryotic genes (9, 10, 21, 24), it is essential to determine the mechanisms underlying differential levels of transcription from varying TATA elements.

To further explore the role of sequence variation in TATA element function, the in vivo and in vitro properties of a panel of substituted elements were characterized (summarized in Table 3). As expected, each of the elements was capable of binding TBP in vitro, albeit to varying extents. We observed that C and G substitutions at the first and sixth position of the TATA element were better tolerated for TBP binding, whereas substitutions in the middle positions typically resulted in decreased TBP binding. Sensitivity of TBP binding to substitutions in the middle bases of the TATA element are likely to arise from an exocyclic NH$_2$ protruding from G in a C-G or G-C base pair. This protruding group has been predicted to disrupt the TBP-DNA interaction by sterically clashing with hydrophobic residues on the concave undersurface of TBP (18). Replacement of A for C or G at position 2 clashes with TBP residue Leu-163. A substitution at position 3 from T to G will artificially increase the effective concentration of these proteins at the promoter element being tested.

The response to Gal4 in the strain containing the TFIIA-TBP or TFIIB-TBP fusion as the sole source of TBP was measured for all 13 elements using the β-galactosidase assay. Incubation in galactose medium for 18 h resulted in significant induction in activity in both the TFIIA-TBP and TFIIB-TBP fusion strains harboring the TATAAA driven reporter (Fig. 7), as expected for a canonical element. Strikingly, the TFIIA-TBP fusion strains containing the CATATAA and TATAAG driven reporters showed a dramatic increase in activity to ~30% of TATAAA activity. This increase in CATATAA and TATAAG activity is specific to these elements because none of the other elements showed any increase in activity in the Toa2-TBP fusion strains. Moreover, the increase is specific for TFIIA because no such increase is observed in the TFIIB-TBP fusion strain. Thus, transcriptional output from elements characterized by an unstable TFIIA-TBP-DNA complex can be specifically enhanced by fusing TFIIA and TBP.
A decrease in the TBP-DNA interaction were typically also compromised in forming the higher order complexes with TFIIA and TFIIIB. Generally speaking, neither TFIIA nor TFIIIB could overcome defects in TBP binding to the TATA variants, with one notable exception. The presence of TFIIIB, but not TFIIA, restored TBP binding to the TATGAA element to the level of complex formed on TATAAA. This result suggests that TFIIIB specifically introduces a conformational change in the TBP-DNA complex that can overcome the steric clash between the exocyclic NH2 group on the guanine and Val-119 of TBP.

Although TATAAG was able to form the initial TBP-DNA, TFIIA-TBP-DNA, and TFIIIB-TBP-DNA complexes to the same extent as the control elements TATAAA and CATAAA, measurements of the stability of the TFIIA-TBP-DNA showed that, like CATAAA (11), TATAAG was defective for the formation of a stable TFIIA-TBP-DNA complex. Experiments measuring the stability of the TBP-DNA complex and TFIIA-TBP-DNA complex on TATAAG revealed that these complexes behaved indistinguishably from TATAAA and CATAAA. Therefore, the difference in stability of protein complexes formed on CATAAA and TATAAG is specific to the higher order complex containing TFIIA and a TFIIA-TBP fusion molecule can restore activity from these two elements in vivo (whereas a TFIIIB-TBP fusion could not). Significantly, the other TATA variants did not show an increase in transcriptional activity in the TFIIA-TBP fusion strain even though some of these elements showed measurable loss of TFIIA-TBP-DNA complex during the competition experiments. Thus, stabilizing the TFIIA-TBP interaction could only restore activity from the elements CATAAA and TATAAG, where the stability of the TFIIA-TBP-DNA complex appears to be the primary mechanistic defect.

CONCLUSIONS

Why study variant TATA elements? It is clear from the analysis of yeast and higher eukaryotic promoter sequences that a majority of genes do not possess TATA elements that match the consensus (8, 15–17, 25, 26). In addition, whole genome approaches reveal that a consensus TATA element does not appear to be a major indicator of in vivo TBP binding or gene expression (27). As such, determining the properties of variant TATA elements with regard to preinitiation complex formation and in vivo gene expression is fundamental for advancing our understanding of the regulation of gene expression by RNA polymerase II.

We have demonstrated that CATAAA and TATAAG share a common mechanism of transcription regulation, namely the altered stability of the TFIIA-TBP-DNA complex. Analysis of native yeast promoters reveals that a substitution at the first position with a C (CATAAA) is the most common substitution at this position with 15% of the elements exhibiting this alteration in yeast (8) and 10% in higher eukaryotic promoters (28). In addition, at the sixth position of the element, the only substitution tolerated in yeast and higher eukaryotes is a G (TATAAG). This suggests that the altered stability of the TFIIA-TBP-DNA complex may provide a useful and widespread mechanism for regulation of gene transcription from non-optimal core promoter elements in vivo. It is important to point out that these elements do support a modest level of response to an activator in vivo (2–4-fold). Two-fold changes in the level of gene expression can have profound effects on cell growth, differentiation, and the response to environmental stresses (29–33). Overall, the data presented here indicate that substitutions at various positions in the TATA sequence result in particular mechanistic penalties, which have the potential to contribute to functionally distinct methods of the regulation of gene expression.

Acknowledgments—We thank Kevin J. Lumb for providing the pYTBP plasmid, Katie Campbell for advice during the TBP purification, and Mary M. Robinson for critical reading of the manuscript.

REFERENCES

1. Nothias, J. Y., Majumder, S., Kaneko, K. J., and DePamphilis, M. L. (1995) J. Biol. Chem. 270, 22077–22080
2. Smale, S. T., and Kadonaga, J. T. (2003) Annu. Rev. Biochem. 72, 449–479
3. Woychik, N. A., and Hampsey, M. (2002) Cell 108, 453–463
4. Hahn, S. (2004) Nat. Struct. Biol. 11, 394–403
5. Chen, W., and Struhl, K. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 2691–2695
6. Singer, V. L., Wobbe, C. R., and Struhl, K. (1990) Genes Dev. 4, 636–645
