Published online 6 October 2006 *Nucleic Acids Research, 2006, Vol. 34, No. 19* ***5585–5593***
doi:10.1093/nar/gkl534

The *Saccharomyces cerevisiae* RNA polymerase III recruitment factor subunits Brf1 and Bdp1 impose a strict sequence preference for the downstream half of the TATA box

Nick D. Tsihlis and Anne Grove*

Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA

Received June 5, 2006; Revised July 7, 2006; Accepted July 10, 2006

ABSTRACT

Association of the TATA-binding protein (TBP) with its cognate site within eukaryotic promoters is key to accurate and efficient transcriptional initiation. To achieve recruitment of *Saccharomyces cerevisiae* RNA polymerase III, TBP is associated with two additional factors, Brf1 and Bdp1, to form the initiation factor TFIIIB. Previous data have suggested that the structure or dynamics of the TBP–DNA complex may be altered upon entry of Brf1 and Bdp1 into the complex. We show here, using the altered specificity TBP mutant TBPm3 and an iterative *in vitro* selection assay, that entry of Brf1 and Bdp1 into the complex imposes a strict sequence preference for the downstream half of the TATA box. Notably, the selected sequence (TGTAAATA) is a perfect match to the TATA box of the RNA polymerase III-transcribed U6 small nuclear RNA (*SNR6*) gene. We suggest that the selected T•A base pair step at the downstream end of the 8 bp TBP site may provide a DNA flexure that promotes TFIIIB-DNA complex formation.

INTRODUCTION

The TATA-binding protein (TBP) plays an integral role in transcription by all three nuclear RNA polymerases, including transcription from promoters without a TATA box (1). In the *Saccharomyces cerevisiae* RNA polymerase (pol) III apparatus, TBP is found in the recruitment factor TFIIIB, along with the TFIIB-related factor, Brf1 and the pol III-specific Bdp1. *In vivo*, TFIIIB is assembled onto the DNA via TFIIIC (2–5), but this requirement can be bypassed *in vitro* if a TATA box is present (6,7), allowing TBP to bind the DNA and nucleate a stepwise TFIIIB assembly requiring Brf1 to bind the TBP–DNA complex before Bdp1 can associate. These TFIIIB complexes are indistinguishable *in vitro* from those assembled by TFIIIC (8). Even when TFIIIB is assembled by TFIIIC, direct interaction of TBP with its cognate site contributes to accurate transcriptional initiation (9).

TBP binds DNA in the minor groove and induces a bend of the order of ~80° (10–15). TBP binds an 8 bp site with the consensus sequence TATAa/tAa/tN, where N is any base (16–18), and an iterative *in vitro* selection performed using *Acanthamoeba* TBP on 84 bp DNA containing 40 randomized positions yielded only four classes of TATA box and a preference for TATATAAG [35 of 54 clones; (19)]. Complex stability measurements using DNase I footprinting and the frequency of a sequence appearing in the clones indicate that TBP is able to differentiate between A:T and T:A base pairs, and that the frequency with which a sequence is selected correlates with complex stability (19).

While an A→G substitution at the second base pair of the TATA box abolishes specific DNA binding by wild-type TBP, a mutant TBP known as TBPm3 was previously isolated from yeast and found to bind to the sequence TGTA as well as to the wild-type TATA box (20). The three mutations (I194F, L205V, and V203T) that confer this altered specificity, are in close proximity in the folded protein and are in a position to interact with the second base pair of the TATA box (20). TBPm3 assembles a stable TFIIIB-DNA complex that is functional for pol III transcription (21).

We are exploiting the specificity of binding of TBPm3 to orient the protein unidirectionally on the DNA and investigate TBP–DNA contacts within the downstream half of the TATA box as a function of Brf1 and Bdp1 association. It has been previously suggested that addition of Brf1 and Bdp1 to the TBP–DNA complex alter its conformation or dynamics: (i) While a missing nucleoside at the downstream end of the TATA box, coinciding with the site of TBP-mediated DNA kinking (base pair −23), significantly enhances complex formation, the TFIIIB complex abrogates this preference, instead preferring missing nucleosides within an

*To whom correspondence should be addressed. Tel: +1 225 578 5148; Fax: +1 225 578 8790; Email: agrove@lsu.edu
Present address:
N. D. Tsihlis, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA

extended region downstream of the TATA box (22), (ii) examination of TFIIIB interacting with a region upstream of the *SUP4* $tRNA^{Tyr}$ gene by photochemical crosslinking showed TBP in proximity to the DNA minor groove, except for contacts to the DNA major groove at base pair −23 of the transcribed strand which were enhanced upon TFIIIB-DNA complex formation (23) and (iii) the structure of a ternary complex composed of TBP, DNA and the primary TBP-binding domain of Brf1 revealed an exceptionally large number of interactions that bury 3230 $Å^2$ of TBP surface area (24).

Here, we use an iterative *in vitro* selection to compare the sequence preference exhibited by TBPm3 and TFIIIB assembled with TBPm3. We show that the sequence preference of TBPm3 is less stringent than that reported for wild-type TBP (19). Notably, entry of Brf1 and Bdp1 into the complex imposes a strict sequence preference for the downstream half of the TATA box that matches the TATA box of the pol III-transcribed U6 small nuclear RNA (*SNR6*) gene.

MATERIALS AND METHODS

Protein purification

Plasmids expressing TBPm3, Brf1 and Bdp1 were generous gifts from E. P. Geiduschek and G. A. Kassavetis, University of California San Diego, CA. Purification of TBPm3 was modified from (25), briefly, plasmid containing the gene encoding TBPm3 was transformed into *Escherichia coli* BL21(DE3)pLysS and grown to $OD_{600} = 0.4$ in LB broth containing 100 μg/ml ampicillin. Protein overexpression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 2 h and the pelleted cells frozen at −80°C. Forty-five ml lysis buffer A [50 mM Tris–HCl (pH 8.0), 0.1 mM EDTA, 5% glycerol, 10 mM β-mercaptoethanol, 300 μg/ml lysozyme and 0.5 mM phenylmethylsulfonyl fluoride (PMSF)] were added to ∼5 g thawed cells and incubated for 1 h on ice. All steps after lysis are carried out at 0–4°C. The lysate was diluted ∼1:1 with 60 ml lysis buffer B [50 mM Tris–HCl (pH 8.0), 1 M NaCl, 5% glycerol, 10 mM β-mercaptoethanol, and 0.5 mM PMSF], $CaCl_2$ added to 0.5 mM, and incubated for 1 h with 10 μl DNase I (10 U/μl). The mixture was dialyzed overnight against 3 l buffer A [50 mM Tris–HCl (pH 8.0), 100 mM KCl, 20% glycerol, 1 mM EDTA, 10 mM β-mercaptoethanol and 0.5 mM PMSF] and loaded on tandem DEAE-heparin columns. TBPm3 was eluted from the heparin column using a linear gradient (100–500 mM KCl) and fractions containing TBPm3 identified by SDS–PAGE. Active fractions were pooled and dialyzed against 2 l buffer A for 2 h prior to loading on a CM-Sepharose column. The protein was eluted as described above. TBPm3 concentration was determined by Coomassie blue-staining of SDS–PAGE gels using BSA as a standard. The activity of TBPm3 was determined by electrophoretic mobility shift assays (EMSA), and the preparation found to be essentially 100% active.

The plasmid containing the gene encoding N- and C-terminally-His-tagged Brf1 was transformed into *E.coli* BL21(DE3)pLysS and grown to $OD_{600} = 0.4$ in LB broth containing 60 μg/ml ampicillin. Protein overexpression was induced with 0.4 mM IPTG for 2 h and the pelleted cells frozen at −80°C. Approximately 5 g of thawed cells were resuspended in 15 mL lysis buffer A supplemented with 1 μg/ml pepstatin and 1 μg/ml leupeptin, lysozyme was added to a final concentration of 300 μg/ml and the suspension was allowed to incubate on ice for 30 min. Tween-20 was added to a final concentration of 0.1%, and the cells sonicated on ice five times for 30 sec. The lysate was diluted ∼1:1 with 20 ml lysis buffer B supplemented with 1 μg/ml pepstatin and 1 μg/ml leupeptin, sonicated on ice five times for 30 seconds, then centrifuged at 20 000× *g* for 1 h at 4°C. The pellet was resuspended in 10 ml Buffer G [50 mM Tris–HCl (pH 8.0), 6 M guanidinium-HCl, 10% glycerol, 7 mM β-mercaptoethanol, 0.5 mM PMSF, 1 μg/ml pepstatin and 1 μg/ml leupeptin], then centrifuged at 20 000× *g* to pellet insoluble material. The supernatant fraction was added to 5 ml nickel-NTA agarose beads equilibrated in buffer G and incubated for 1 h at 4°C. The beads were washed three times for 15 min at 4°C with 10 ml buffer G. After harvesting the beads by centrifugation, the protein was eluted by a pH gradient (6.7, 6.5, 5.9, 5.7, 5.5, 5.1 and 4.7), accomplished via successive 15 min washes at 4°C with 10 ml buffer B (7 M urea, 7 mM β-mercaptoethanol, 0.5 mM PMSF, 1 μg/ml pepstatin, and 1 μg/ml leupeptin and 100 mM sodium phosphate at appropriate pH), and fractions containing the double His-tagged Brf1 determined via electrophoresis on a 12% SDS–PAGE gel. The active fractions were pooled, the pH adjusted to 7.9 with 1.5 M Tris–HCl (pH 8.7), and urea removed by sequential dialysis for 2 h against 500 ml buffer C-500 [20 mM HEPES (pH 7.8), 10% glycerol, 7 mM $MgCl_2$, 500 mM NaCl, 10 mM β-mercaptoethanol, 0.01% Tween-20, 0.5 mM PMSF and 10 μM $ZnSO_4$] containing 3.0, 1.5, 0.75 and 0 M urea, respectively (8). The preparation was judged by Coomassie blue-staining of SDS–PAGE gels to be without contamination by truncated variants. Protein concentration was determined by Coomassie blue-staining of SDS–PAGE gels using BSA as a standard. EMSA measuring assembly of TFIIIB suggests that the preparation is ∼15% active.

Bdp1 was overexpressed as described for Brf1 and purification procedures carried out at 0–4°C. Cells were resuspended in buffer W (50 mM potassium phosphate buffer pH 7.0, 350 mM KCl, 5% glycerol, 4 mM β-mercaptoethanol and 0.2 mM PMSF). Lysozyme was added to 0.5 mg/ml followed by incubation for 1 h on ice. Following addition of 10% Triton X-100 to a final concentration of 0.5% (v/v), polymin P was added dropwise from a 13% solution to a final concentration of 0.5%. Cell debris and precipitates were removed by centrifugation for 10 min at 6000× *g*. The supernatant was incubated with Talon beads (BD Biosciences) for 1 h, washed twice with buffer W, and the N-terminally His_6-tagged Bdp1 eluted batchwise by successive 15 min incubations with buffer W supplemented with increasing concentrations of imidazole (10, 25, 50, 75, 100 and 150 mM). Fractions containing Bdp1 were pooled, diluted 1:2 with buffer W and loaded on CM-Sepharose equilibrated in buffer W. Bdp1 was eluted with a linear gradient from 350 mM to 1 M KCl in buffer W. Protein concentration was determined by Coomassie blue-staining of SDS–PAGE gels using BSA as a standard. EMSA indicates that the preparation is at least 25% active. TFIIIB assembled

with the proteins used for selection studies is transcriptionally active (data not shown).

TBPm3–DNA complex formation using EMSA

Oligonucleotides used to generate duplex DNA containing 8 bp TATA- (5′-CGT GAC TAC **TAT AAA TA**G ATG ATC CG-3′) or TGTA-boxes (5′-CGT GAC TAC **TGT AAA TA**G ATG ATC CG-3′) were purified on denaturing polyacrylamide gels. For EMSA, the top strand was 5′ end-labeled using T4 polynucleotide kinase and [γ-^{32}P]ATP, and annealed to the bottom strand by heating to 90°C, followed by slow cooling to room temperature.

Reactions for kinetics assays contained 44 mM Tris (pH 8.0), 8.4 mM NaHEPES (pH 7.8), 50.5 mM NaCl, 7 mM $MgCl_2$, 1 mM EDTA, 8% (v/v) glycerol, 3 mM DTT, 4 mM β-mercaptoethanol and 84 μg/ml BSA. Samples were subjected to electrophoresis on native 10% polyacrylamide gels and in buffer containing 0.5× TBE (45 mM Tris-borate, pH 8.0 and 1 mM EDTA) and 2.5 mM $MgCl_2$.

To determine the rate of complex dissociation during electrophoresis (k_{diss}), 200 fmol TBPm3 and 50 fmol DNA were incubated for 55 min, 400 ng competitor DNA added [poly(dA–dT):poly(dA–dT)], and subjected to electrophoresis for time t. The gels were dried, exposed to a phosphorimaging screen, and the data quantitated using ImageQuant 1.1. Data were fitted to $F_{obs} = F * \exp(-k_{diss}t)$, where F_{obs} is the observed fractional complex, F is the fractional complex present at $t = 0$, k_{diss} is the rate of dissociation on the gel, and t is the time of electrophoresis (26).

For determination of the off-rate in solution, 750 fmol of DNA and 3000 fmol TBPm3 was incubated at room temperature for 1 h, and aliquots loaded on the gel at time t after addition of 6000 ng poly(dA–dT):poly(dA–dT). Data were corrected by $F_{corr} = F/\exp(-k_{diss}t)$, where F_{corr} is the fractional complex corrected for dissociation during electrophoresis, F is the observed fractional complex and t is the time of electrophoresis. The corrected data were fitted to $F_{corr} = F_0 * \exp(-k_{off}t)$, where F_0 is the fractional complex present before addition of competitor, k_{off} is the off-rate in solution, and t is time after addition of poly(dA–dT):poly(dA–dT).

The on-rate in solution was determined for a protein concentration range of 20–80 nM. Protein and 50 fmol DNA were incubated for time t and loaded on the gel immediately after addition of poly(dA–dT):poly(dA–dT). The observed fractional complex was corrected for dissociation during electrophoresis as described above and fitted to $F_{corr} = F_{final}[1 - \exp(-k_{obs}t)]$, where F_{final} is the calculated fractional complex present at completion of the reaction and k_{obs} the apparent first-order rate constant. The reciprocal of the slope of a plot of $1/k_{obs}$ versus 1/[protein] yielded the second-order rate constant, k_{on} (26). All rate constants represent the mean of at least three experiments.

Determination of TBPm3 and TFIIIBm3 sequence preference by iterative *in vitro* selection

The oligonucleotide (5′-CGC TGC AAT CTC TTT TTC AAT TGC TCC GGA C**TG TAA** NNN NGC GGT CCC TTA CTC TTT CCT CAA CAA TTA ACG GCC C-3′, mutant TATA box underlined and bold) was purified on a denaturing 5% polyacrylamide gel, and amplified by PCR using *Taq* polymerase and 40 pmol of primers PSXB (5′-GGG CCG TTA ATT GTT GAG-3′) and PSXT (5′-CGC TGC AAT CTC TTT TTC-3′). Reaction conditions included buffer supplied by the manufacturer containing 2.0 mM $MgCl_2$ and 250 μM dNTP. The starting pool of oligonucleotides easily contains every possible sequence (4^4 or 256 sequences). The double-stranded product was 5′ end-labeled using T4 polynucleotide kinase and [γ-^{32}P]ATP and at least 40 ng was incubated with 400 fmol TBPm3 to yield no more than ~10% complex in early rounds (this fraction should yield a consensus sequence from the 256 possible sequences within 4–5 rounds of selection) for 1 h in the buffer listed above, except with 150 mM NaCl. After addition of 800 ng poly(dA–dT):poly(dA–dT), the reaction was loaded onto a native 10% polyacrylamide gel with the power on and subjected to electrophoresis at 175 V for 1 h. The gel was exposed to a phosphorimaging screen, the TBPm3–DNA complex excised from the gel, and the DNA passively eluted overnight in 1 ml elution buffer [20 mM Tris–HCl (pH 8.0), 1 M LiCl, 0.2 mM EDTA, 0.2% SDS] with rotation. The recovered DNA was amplified by PCR as described above, the PCR product purified on a native 7% polyacrylamide gel, radioactively labeled, and used as template for the next round of selection. After 10 rounds of selection, the DNA was cloned into the pCR T7/NT-TOPO vector (Invitrogen) and transformed into *E.coli* TOP10. Sequences containing the TGTAA box were aligned with ClustalX (27).

For the TFIIIBm3 selection, the selection was performed using the same conditions as for TBPm3, except that 40 ng labeled DNA was incubated with 120 fmol TBPm3, 520 fmol of total Brf1 and 1200 fmol of total Bdp1 for 1 h, with 100 mM NaCl. Active Brf1 was chosen as the limiting component to avoid trapping all existing TBPm3–DNA complex (28). Poly(dA–dT):poly(dA–dT) (240 ng) was added, and the reaction was loaded onto a native 4% polyacrylamide gel with the power on and subjected to electrophoresis at 175 V for 3 h. The recovered DNA was amplified by PCR, the PCR product purified on a native 10% polyacrylamide gel, radioactively labeled, and used as template for the next round of selection. After 10 rounds of selection, the DNA was cloned into the pCR4-TOPO vector (Invitrogen) and transformed into *E.coli* TOP10. Sequenced DNA was aligned using ClustalX (27).

To confirm complex formation by TBPm3, EMSA was performed as described above using 26 bp constructs, except that the sequence was modified to represent sequences selected by TBPm3.

Two-dimensional methidiumpropyl-EDTA (MPE)-Fe(II) footprinting

Oligonucleotides were purchased and purified on 5% denaturing polyacrylamide gels. The top strand of the 76 bp DNA probe (5′-CGC TGC AAT CTC TTT TTC AAT TGC TCC GGA CTG TAA **ATT G**GC GGT CCC TTA CTC TTT CCT CAA CAA TTA ACG GCC C-3′) was 5′ end-labeled using T4 polynucleotide kinase and [γ^{32}-P]ATP, and annealed to the unlabeled complementary strand. One hundred twenty-five fmol of 76 bp DNA was incubated with or without 500 fmol TBPm3 at room temperature for 1 h using the buffer conditions described above, except

with 3 mM $MgCl_2$. After addition of 1 μg poly(dA–dT):poly(dA–dT), 1 μl of 10 mM sodium ascorbate and 4 μl freshly prepared 25 μM Fe-MPE were added, incubated for 1 min and the reaction was stopped by loading onto a native 10% polyacrylamide gel with the power on and subjected to electrophoresis at 175 V for 1 h. Free DNA and TBPm3–DNA complex were excised from the gel, and the DNA eluted and purified, as described above.

Samples were resuspended in formamide loading buffer and heated at 95°C for 2 min prior to loading on a 10% polyacrylamide sequencing gel. The gel was run at 35 W in 1× TBE for ~4 h, and dried. The gel was exposed to a phosphorimaging screen, and the gel image quantitated in ImageQuant 1.1.

RESULTS

Characterization of TBPm3

TBPm3 does not fully substitute for wild-type TBP *in vivo*, as evidenced for example by the slower growth phenotype of yeast strains carrying TBPm3 as the only TBP variant (20). As a basis for comparison of sequence preferences exhibited by TBPm3 and TFIIIB assembled with TBPm3 to those reported for wild-type TBP, we therefore first determined the affinity of TBPm3 for DNA constructs carrying either an 8 bp TATA box (the U6 TATA box) or the corresponding TGTA box. To calculate the equilibrium binding constant, rates of association and dissociation in solution were determined using EMSA (26,29) (Table 1). As shown in Figure 1, the TBPm3–DNA complex is quite stable in solution and decays with first-order kinetics with $t_{1/2}$ of 72 and 61 min, respectively, for DNA containing either the TATA or TGTA box. The observed rate of dissociation is comparable with that observed for wild-type TBP (26,29). Association of TBPm3 with DNA is detectable after 10 s, and a gradual increase in complex formation is observed after longer incubation times. The second-order rate constants for association of TBPm3 with DNA containing either the TATA or TGTA box are comparable (5.3×10^5 and 3.1×10^5 $M^{-1}s^{-1}$, respectively, Table 1 and Figure 2) and well within the range of values reported for association of wild-type TBP with various DNA substrates, using a variety of techniques (26,29–35). The calculated equilibrium dissociation constant for TBPm3 binding to the TATA or TGTA probe is 0.3 and 0.6 nM, respectively.

Determination of TBPm3 sequence preference for the downstream half of the TGTA box

The ability of TBPm3 to bind the TGTA box unidirectionally was exploited to perform an iterative *in vitro* selection on a 76 bp DNA construct in which four bases at the downstream end of the TGTA box were randomized; this DNA construct derives from a modified $tRNA^{Tyr}$ gene in which a 6 bp TATA box was embedded in a G+C-rich surrounding sequence (9). The randomized region was selected to coincide with sites at which modulation by Brf1 and Bdp1 may be expected and includes positions 6, 7 and 8 of the TBP site and one base pair downstream (22,23). Positions 1–5 of the TBP site were retained to ensure unidirectional binding, and inclusion of 1 bp downstream of the 8 bp TBP site in the randomized segment was chosen as TBP has been seen also to exhibit a sequence preference at this position (19). For stringency of selection, the concentration of NaCl was raised to 150 mM

Table 1. Rates of complex formation and dissociation

TATA	TGTA
k_{diss}: $7.0 \pm 2.9 \times 10^{-3}$ min^{-1}	k_{diss}: $6.6 \pm 1.3 \times 10^{-3}$ min^{-1}
k_{off}: 1.6×10^{-4} s^{-1}	k_{off}: 1.9×10^{-4} s^{-1}
k_{on}: 5.3×10^5 M^{-1} s^{-1}	k_{on}: 3.1×10^5 M^{-1} s^{-1}
K_d: 3.0×10^{-10} M	K_d: 6.0×10^{-10} M
$t_{1/2}$ (solution): 72 min	$t_{1/2}$ (solution): 61 min

The left column represents data for 26 bp DNA containing the U6 TATA box. Data in the right-hand column correspond to the TGTA sequence. Errors for the off-rates are 5% for TGTA and 11% for TATA DNA, errors for the on-rates are 20% for TGTA and 27% for TATA DNA.

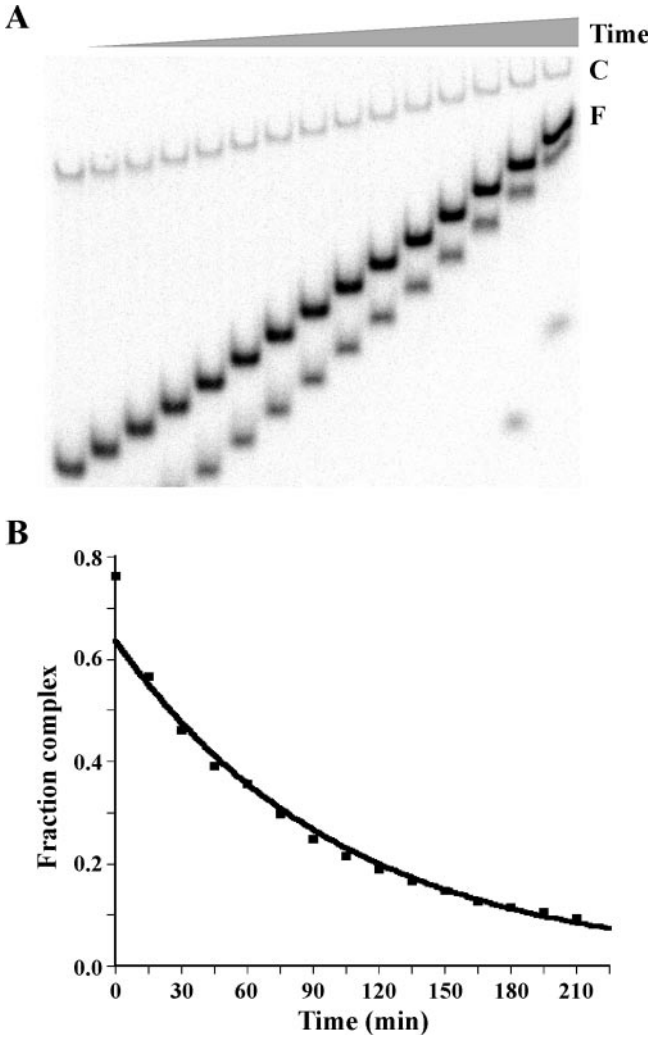

Figure 1. Determination of the dissociation rate constant, k_{off}. (**A**) TATA DNA was incubated with TBPm3 for 1 h, and aliquots loaded on the gel at time *t* after addition of poly(dA–dT):poly(dA–dT). *C* indicates the TBPm3–DNA complex and *F* indicates the free dsDNA. The faint band below *F* is ssDNA. (**B**) Rate of complex dissociation. The fraction of complex corrected for dissociation during electrophoresis is shown as a function of time after addition of competitor.

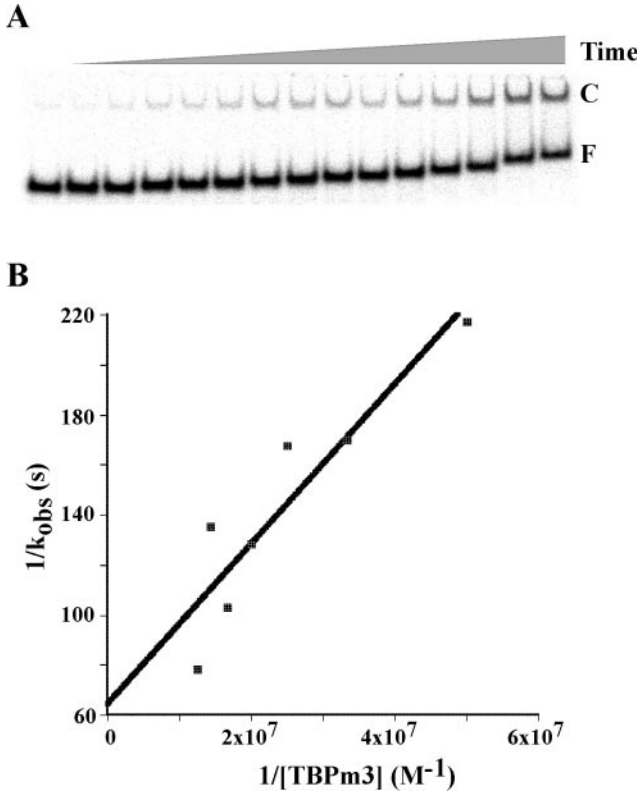

Figure 2. Determination of the association rate constant, k_{on}. (**A**) TBPm3 (50 nM) and TGTA DNA were incubated for time t (10 s to 10 min) and loaded on the gel immediately after addition of competitor to determine the apparent first-order rate constant, k_{obs}. *C* indicates the TBPm3–DNA complex and *F* indicates the free dsDNA. (**B**) The kinetics of TBPm3–DNA complex formation were determined at the indicated protein concentrations. The reciprocal of k_{obs} is plotted as a function of the reciprocal of the TBPm3 concentration.

Table 2. Frequency of occurrence of individual bases at each of the four randomized positions after selection by TBPm3

	N1	N2	N3	N4
A	23	15	15	7
C	7	6	9	8
G	20	15	17	32
T	16	30	25	19
Favored	a/g	T	T	G

during incubation of TBPm3 with DNA. With only 256 possible sequences, a consensus should be reached within 4–5 rounds of selection. However, the sequence preference exhibited by TBPm3 was only modest after five rounds of selection (data not shown), so we elected to continue the selection for another five rounds. After 10 rounds of selection, the selected pool of DNA was again sequenced and a favored sequence determined for the four randomized bases (Table 2). Of the sequenced clones, 66 contained the original TGTAA sequence and were used to determine the consensus for this selection. A total of 29 clones contained a TATAA box, suggesting its generation as a result of errors introduced during PCR; these sequences were not included in the alignment to avoid potential introduction of sequences arising from TBPm3 binding in the reverse orientation. A total of 17 clones contained a sequence comprised of a series of GTG repeats, with only the regions complementary to the primer sequences constant. The remainder of the clones contained sequence with no match to either of the above categories, such as other alterations to original TGTAA sequence. Alignment of the 66 TGTAA-containing sequences still showed a surprisingly modest sequence preference for each of the randomized positions. Whereas a C is generally disfavored at every position, position N1, corresponding to the sixth base pair of the TBP site, shows essentially only selection against C. Positions 7 and 8 of the TBP site (N2 and N3) reveal a modest preference for T, while a G is preferred at position N4. This is in contrast to bases selected by wild-type TBP, for which a G at positions equivalent to N1 and N2 was not observed (Figure 3). Apparently, TBPm3 exhibits a less stringent sequence preference compared with wild-type TBP.

The results of the TBPm3 selection were verified by EMSA and MPE-Fe(II) footprinting on a DNA probe representing the most frequently selected bases at each position, TGTAAATTG (note that this sequence represents the most frequently occurring bases at each randomized position, but was not found among the selected clones). TBPm3 was seen to bind to 26 bp DNA containing this sequence, while disfavored sequences, containing for example a C in position N1 (TGTAACTGG) yield barely detectable complex formation (Figure 4 and data not shown). MPE-Fe(II) footprinting on the 76 bp DNA containing the favored sequence indicates that TBPm3 is binding at the TGTA box, despite the fact that this sequence was not found among the selected clones, as seen by the partial protection from cleavage at positions −28 to −23 (Figure 5), where the first T of the TGTA box is designated −30. Enhanced cleavage was observed at base pair −19, −18 and −31 consistent with that observed for wild-type TBP at these positions (22,29).

When the selection was performed using TFIIIB assembled with TBPm3, a distinct sequence preference emerged. A total of 25 clones containing the TGTAA sequence were aligned to determine the consensus for this selection. As for the TBPm3 selection, clones containing the TATAA sequence (25 clones) were excluded from the alignment, as were 13 clones containing the series of GTG repeats. While the occurrence of the GTG-repeat sequences in both selections is curious, we did not pursue this observation further. The remainder of the clones contained sequence with no match to the above categories, including other alterations to the 5′ half of the TGTA box. Alignment of the 25 TGTAA-containing sequences (Table 3) showed a much stronger sequence preference compared with TBPm3 alone, despite reaction conditions that should have allowed less stringent binding (100 mM NaCl versus 150 mM for the TBPm3 selection; Figure 6). Notably, the selected consensus sequence (TGTAAATAG) is a perfect match to the 8 bp U6 TATA box (TATAAATA).

DISCUSSION

TBPm3–DNA complex formation

TBPm3 dissociates from its DNA site with first-order kinetics and exhibits second-order kinetics of association, as reported

AATTACTCCGGAT TGTAA**ATTA** CCGATCCTTTAC-CTTTC
AATTACTCCGGAT TGTAA**ATTA** CCGATCCTTTAC-CTTTC
AATT-GTCCGGAT TGTAA**ATTA** GCGGTCATTTACTGCTTT
AATCACTCGGGAG TGTAA**ATGG** GCAGTCCT-TACTCTCTC
AATCACTCGGGAG TGTAA**ATGG** GCAGTCCT-TACTCTCTC
AATTGCTCCGGAC TGTAA**ATGG** GCGGTCCTTTACTCTTTC
AATTGCTCGGGAT TGTAA**ATGT** GCGGCCCATTATTCTTCC
AATGGCTCCAGAC TGTAA**ATAG** GCATTCCCTTACCTTTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATTGCTCCGGAC TGTAA**ATAG** GCGGTCCCTTACTCTTTC
AATGGCTCCAGAC TGTAA**ATAG** GCATTCACTTACTGCTTC
ATTGGCTCCGCAC TGTAA**ATAC** GTGGTTGTGTTTCGCGGG
ATT-CCTGCAGAC TGTAA**ATAC** GGTTGGTTGGGGGGTGGG
ATTG-CTCTGAAC TGTAA**ATAC** GTGGTTGTGTTTCGCGGG
AATTACTCCGAAC TGTAA**ATAC** GCGGTTGTGTTTCGCGGG
AATTGCTCCGCAC TGTAA**ATCT** GCGGCCACTTACTCTTTC
A-TTG-TCCAGAC TGTAA**AAGA** GCGGTCCCTGACTATTTC
AACTGCCCCGGAC TGTAA**AAAT** TCGGTCCTTTATTCTTTC
AATTGGTCCGGAC TGTAA**AGTG** GCGGTCCCTTACTCTTTG
AATTGCTCGGGAC TGTAA**AGGG** CCGGCCCCTTACCGTTTC
AATCGTTCCGGTC TGTAA**AGTC** GTGGCCACTTACTCTTTC
AACTACCCCGGGC TGTAA**ACAG** GCGATCACTTACCTTTAC
A-TTGCTCGGGAC TGTAA**ACTA** ACGGCCACTTAATCTTCC
AATTGCTCCGGTC TGTAA**GTTG** GCGATCCCTTAC-CTTTC
AA-TGTTCCGAGC TGTAA**GTTG** GCGATTACTTACTTCTTC
AACTGCTCCGGAC TGTAA**GTTT** GCGGCCCCTTACCTTTTC
AATCGCTCCAGAC TGTAA**GTTT** GCGGTCAGTTACCTTTTC
AATCGCTCCAGAC TGTAA**GTTT** GTGGTCAGTTACCTTTTC
AATTGCTCCGGAC TGTAA**GTGT** GCGTCTC-TTACTCCTTC
AATTGCTCCGGAC TGTAA**GTGT** GCGGC-C-TTAC-CTTTC
AATTCCTCGGGAC TGTAA**GTGT** GCGTTCACTTACTCTTTC
A-TTGTTCCGGAC TGTAA**GATG** GCTGTCTCTTACTCTTCC
ATTGACTCGGGAT TGTAA**GATT** GGGATCCCTTACCCTTTC
AATTGCTCCGGAC TGTAA**GAAG** GCGGTCTCTTACTCTTTC
AATTGCTCCGGAC TGTAA**GAAG** GCGGTCTCTTACTCTTTC
AAGTGCTGCGAAC TGTAA**GGTG** GGGATCTCTTATACTTTC
AAGTGCTGCGAAC TGTAA**GGTG** GGGATCTCTTATACTCTC
AAGTGCTGCGAAC TGTAA**GGTG** GGGATCTCTTATACTTTC
AAGTGCTGCGAAC TGTAA**GGTG** GGGATCTCTTATACCTTC
AAGTGCTGCGAAC TGTAA**GGTG** GGGATCTCTTATACTTTC
AATCGCTCCGGAC TGTAA**GGAT** GCGGCCCTTTACCTTTTC
AATGGCTCCGGCT TGTAA**GGGT** GCGACCACTTACACTCTC
AATTGCTCCGGAC TGTAA**GCGG** GCGATCCTTATCTATTTC
AATTCCTCCAGTC TGTAA**TTCG** GCGGCCCCTTATTTCTTC
AATTCCTCCAGTC TGTAA**TTCG** GCGGCCCCTTATTTCTTC
AATTGCTCCAGTC TGTAA**TTCG** GCGGCCCCTTATTTCTTC
AATGGCTCCGGAC TGTAA**TATG** GCGGCCACTTACTCTTTC
ATTGGCTCCGGAC TGTAA**TATG** GCGGTCACTTACTCTTTC
ATTGGCTCCGGAC TGTAA**TATG** GCGGTCACTTACTCTTTC
AATCACTCGGGAC TGTAA**TATT** GCGTTCCCTTACTCTTTC
AGTGGCTCCGGAC TGTAA**TAGC** GCGGTACCTTACTCTTTC
AGTGGCTCCGGAC TGTAA**TACT** GCGGTACCTTAC-CTTTC
AGTGGCTCCGGAC TGTAA**TACT** GCGGTACCTTAC-CTTTC
AATTACTCCGAAC TGTAA**TGTT** GCGGTCCTTTACTTCTTC
AATTGCTCCGGAC TGTAA**TGGC** GCGGTCCCGTACCTTTTC
AATGGCCCCGGAC TGTAA**TGAT** GCGTTCCCTTACCTTTGC
ACTTG--TCGGAC TGTAA**TGAT** GCGATCTCTTACTGTTTC
AATTGCTCCAGAC TGTAA**TCTG** GCGGTCCCTTACCTTTTC
AATTGCTCCAGAC TGTAA**TCTG** GCGGTCCCTTACCTTTTC
AATTGCTCCAGAC TGTAA**CTGT** GCGGTCCATTACCCCTCT
AATGGCTCCGGAC TGTAA**CTCA** GCGGCCTCTTACTCTTTC
AATGGCTCCGGAC TGTAA**CTCA** GCGGCCTCTTACTCTTTC
AATTGCTCCAGAC TGTAA**CACG** GCGATTCCTTACTTCTTC
AATCGCTCCAGAC TGTAA**CAGG** CCAGTCCTTTACTCTCTC
AATTGCTCCGGAC TGTAA**CGGC** GCGGCCCCTTATTTCTTC
AATGGTTCCGGAC TGTAA**CCGG** GCAGTCCTTAACTCT-TC

Figure 3. Alignment of sequences selected by TBPm3. Only sequences retaining the original TGTAA sequence are shown. Bases corresponding to randomized positions are in boldface.

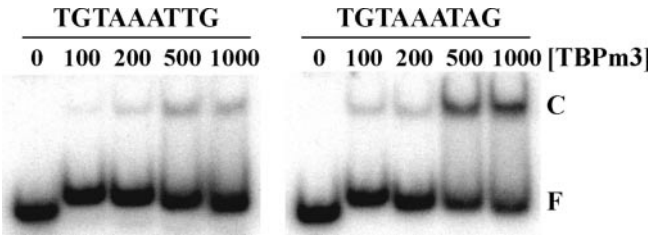

Figure 4. TBPm3 binds to DNA representing the selected sequence. Of each 26 bp 50 fmol DNA was incubated with 0, 100, 200, 500 and 1000 fmol TBPm3. Left panel, selected sequence, right panel, TGTA probe used for affinity determinations. *C* indicates the TBPm3–DNA complex and *F* indicates the free dsDNA.

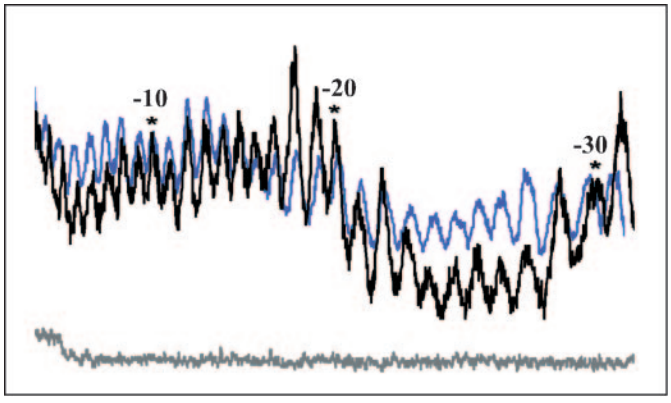

Figure 5. MPE-Fe(II) 2D footprinting confirms binding of TBPm3 at the TGTA box. Densitometry profiles of 76 bp DNA containing the favored sequence (TGTAAATTG) incubated with (black line) and without (blue line) TBPm3 show protection at the TGTA box. Numbering is based on the start site of transcription (+1). Gray line represents uncut DNA.

Table 3. Frequency of occurrence of individual bases at each of the four randomized positions after selection by TFIIIBm3

	N1	N2	N3	N4
A	23	1	21	1
C	0	1	0	2
G	1	2	2	19
T	1	21	2	3
Favored	A	T	A	G

for wild-type TBP (26,29–35). As also seen for wild-type TBP under comparable experimental conditions, rate determinations do not indicate any contribution from a competing TBPm3 monomer–dimer equilibrium (29,36). Rates of association with either DNA probe are within the range reported for wild-type TBP, while the rate of dissociation is slower ($t_{1/2}$ of 61 and 72 min) compared with ~10 min for wild-type TBP using DNA containing the 8 bp U6 TATA box (29). This difference may be owing to the lower [NaCl] used here, as shown by the enhanced rate of dissociation of wild-type TBP that accompanies an increase in [KCl] from 60 to 80 mM [$t_{1/2}$ 100 versus 65 min using the AdML promoter TATA sequence; (26)]. In addition, more stable complex formation may be the consequence of sequence flanking the 8 bp TATA box [the A immediately downstream of the U6 TATA box used in previous assays (29) was replaced with a G in the constructs used here]; while TBPm3 dissociates from TATAAATAG with $t_{1/2} = 72$ min (Table 1), $t_{1/2}$ for dissociation from TATAAATAA is 53 min (data not shown). Sequence flanking the TATA box

AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
A-TGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCACTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** ACATCCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
A-TGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AACGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGACCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
AATGGCCCCAGAC TGTAA**ATAG** GCATTCCCTTTCTGCTTC
A-TTGCTCTGAAC TGTAA**ATAC** GTGGTTGTGTTTCGCGGG
AATTGCTCCGGAC TGTAA**AGTT** GCGCTCCCTTACTCTTTC
AATTGCTCCGGAT TGTAA**AGGT** GCGGTCCCTTACTCTTTC
CATTGCTCCGGAC TGTAA**ACAA** GCGGTCACTTACTCTTTC
AATTGCTCCGGAC TGTAA**GTGT** GCGATCCCTTACTCTTTC
AAATGCTCCGGAC TGTAA**TATC** GCGGTCCCTTATTCTTTC

Figure 6. Alignment of sequences selected by TFIIIB. Only sequences retaining the original TGTAA sequence are shown. Bases corresponding to randomized positions are in boldface.

has been previously shown to modulate kinetics of TBP dissociation, in particular for TBP sites characterized by alternating A-T base pairs (35,37).

TBPm3 binds to the TGTA and TATA probes with comparable affinity, but we note that the modestly higher affinity for the TATA-containing DNA ($K_d = 0.3$ nM versus 0.6 nM for TGTA-containing DNA) is consistent with the identification of numerous TATA-containing sequences in the *in vitro* selections (Figures 3 and 6). The basis for this difference in affinity may be the increased flexibility of the T•A step relative to the T•G step (38,39). As for wild-type TBP, the rate of association of TBPm3 with DNA is orders of magnitude slower than the diffusion limit; for wild-type TBP, the rate of association appears not to be affected by flexure at the sites of DNA kinking, whereas complex stability is (29). Consistent with this observation, rates of association of TBPm3 with either TATA- or TGTA-containing DNA are equivalent.

Sequence preference of TBPm3. The orientation of TBP on the TATA box is such that the C-proximal TBP domain interacts with the 5′ half of the TATA box, while the N-proximal domain contacts the less-conserved 3′ half-site (11–13,40). Sequence specificity at the upstream half of the TATA box has been suggested to be in part imposed by the presence of a proline (Pro191) that would disallow any base other than a T at the 5′ end of the TATA box owing to steric clashes with other bases (13). The equivalent residue in the N-proximal TBP repeat is alanine (Ala100) which imposes no such steric constraints. The modest orientational preference of TBP observed *in vitro* has been suggested to derive also from differential DNA flexure at the two sites of kinking (29). In the preinitiation complex (PIC), however, the orientation of TBP is largely determined by interaction with other transcription factors (21,29,41).

For TBPm3, three substitutions create a binding pocket that can accommodate G at position two of the TATA box. TBPm3 exhibits an only modest sequence preference for the last four bases of the TGTA box, with C generally disfavored at every position. While A→T and T→A transversions cause little change to the chemical environment of the DNA minor groove, the introduction of GC or CG base pairs results in the exocyclic amino group of G protruding into the minor groove. For wild-type TBP, cavities in the interface between TBP and the DNA minor groove can be seen to accommodate a G in positions 3 and 6 of the TATA box (40). The frequent occurrence of a G at position N1 (position 6 of the TBP site) was somewhat unexpected, but this portion of the helix is flattened and unwound in the wild-type TBP–DNA co-crystal structure, and there may likewise not be steric clashes between TBPm3 and the DNA. The widening of the minor groove that accompanies bending into the major groove is more difficult with GC base pairs, hence the more easily deformable TA sequence is preferred by wild-type TBP. Perhaps TBPm3 features additional contacts that may support bending of more rigid sequences.

We note also that the bases most frequently selected at each position do not occur together. In its association with DNA, TBP introduces a significant bend at both ends of the TATA box. The energetically most favorable bending of *B*-DNA occurs by compression of the major groove with concomitant opening of the opposing minor groove; consequently, TBP generally targets A+T-rich regions that feature a greater range of minor groove widths. An exception is poly(dA) runs that exhibit local structural rigidity, as seen by the interlocking of major groove methyl groups of consecutive thymines (13,38,39,42). In an A•T base pair step, the stacking of the methyl group of thymine against the adjacent adenine is likewise extensive. Consequently, TBPm3 may select against the sequence TGTAAATTG owing to its inherent stiffness, even though each base is the most frequently selected at its respective position.

The sequence most frequently selected by wild-type TBP, TATATAA is followed by a G to complete the 8 bp TBP site, with a G or C found at the position immediately downstream (19). This sequence is selected against in our assay as the first five bases of the TGTA box (TGTAA) were held constant, thus a T in position five of the TATA box could only have arisen as a result of errors during PCR. Notably, of all the selected sequences, only one featured the sequence TGTAT, suggesting that it is not favored by TBPm3 (unlike the sequence TATAA, which occurred in 29 of the clones, despite position two of the TBP site also not corresponding to a randomized position). Selection by wild-type TBP for a sequence that includes an A at position five of the TATA box was followed by the sequence A-T-A, generating the U6 TATA box TATAAATA, with a C preferred at the position immediately downstream of this 8 bp TBP site (4 of 54 clones; (19)). Eight base pair alternating TA sequences were generally followed by a G or C (12 of 54 clones). Accordingly, the presence of a G following the 8 bp TBP site preferred by TBPm3 is consistent with the preferred base following an 8 bp A+T-containing sequences selected by wild-type TBP (19,37). Since sequence flanking the 8 bp TBP site affects complex stability but not the rate at which TBP associates with the TATA box, flanking the A+T-rich TBP site with G+C-rich sequence may create border effects that stabilize bound TBP (35,37).

Brf1 and Bdp1 impose a strict sequence preference for the downstream half of the TATA box

The sequence preference of TFIIIB assembled with TBPm3 for the downstream half of the TGTA box differs significantly from that exhibited by TBPm3 alone. While TBPm3 mainly discriminates against C in positions N1–N3, entry of Brf1 and Bdp1 into the complex imposes a strict preference for the sequence A-T-A. In both selections, a G at position N4 is preferred, although only modestly so for TBPm3. Comparable with the TBPm3 selection, no sequences occur in the TFIIIB selection with the sequence TGTAT. We also discount the possibility that TFIIIB may reverse orientation, as a C is strongly disfavored at position N2. Accordingly, the sequence selected by TFIIIB matches that of the native U6 TATA box, except that a G immediately downstream of the 8 bp TATA box is seen in preference to the naturally occurring A.

It was previously shown that stability of TBP on a 6 bp TATA box, which is suboptimal for TFIIIB assembly, is comparable with that of an 8 bp TATA box, which efficiently supports assembly of TFIIIB (29). The significant difference between TATA box sequences must therefore be structural or dynamic adaptations to interaction with Brf1 and Bdp1. In general, the DNA bending that occurs upon association with TBP brings flanking DNA segments closer together to facilitate contacts with other transcription factors that make up the PIC (24,43–45), and sequences that promote a disposition of DNA flanking the TBP-mediated DNA bends in a direction consistent with association of Brf1 and Bdp1 may be preferred. Indeed, analysis of TBP in complex with several divergent TATA sequences reveals comparable structures, yet only some are permissible for PIC formation; base pair changes may well be tolerated in terms of binding to TBP, but may negatively affect recruitment of other transcription factors (40).

The efficiency with which the TBP–TATA complex promotes transcriptional activity depends on the sequence of the TATA box, including A-T transversions that do not alter functional groups present in the DNA minor groove. Presumably, TBP depends significantly on recognition of inherent flexibility of the TATA box, and such differences may also affect PIC assembly (40,46). For example, molecular dynamics simulations of different TATA variants suggest that DNA flexibility is correlated with transcriptional activity by RNA pol II (47). Correlating molecular dynamics simulations of TBP–TATA complexes involving different TATA sequences with reported transcriptional activity by pol II further suggests that optimal pol II activity occurs on DNA that allows the two domains of TBP to rotate relative to each other and that allows the H2 helix of TBP to assume an optimal disposition to interact with factors that bind both TBP domains (such as Brf1). In contrast, low activity DNA sequences appear to promote movement of the H1 helix of TBP and to involve conformational changes in the DNA (48).

TBP introduces roll deformations at either end of the TATA box (11–13). The T•A base pair step is easily deformable owing to its large range of allowable roll angles and is often found in DNA sequences requiring a sharp bend (38,39,49). Indeed, roll deformations at the downstream kink of TATA DNA in complex with TBP vary from ~30° for A•G steps to >45° for T•A steps (49). A unique feature of the U6 TATA box sequence identified in our selections with TFIIIB is the presence of a T•A step at the downstream end of the 8 bp TBP site. While this sequence is not strongly favored by either wild-type TBP or TBPm3 alone, it clearly promotes formation of the TFIIIB-DNA complex. Consistent with this interpretation, *in vitro* transcription with *Drosophila* nuclear extract indicated that while pol II utilizes the TATA box TATAAAAA in the forward direction, pol III reverses orientation (50). We suggest that the unique feature of the selected sequence is a flexibility at the downstream end of the 8 bp TATA box that promotes Brf1 and Bdp1 binding and the associated DNA deformation downstream of the TATA box (22).

ACKNOWLEDGEMENTS

The authors are grateful to E. P. Geiduschek and G. A. Kassavetis for providing expression vectors. Supported in part by the National Science Foundation (MCB-0414875 to A.G.) Funding to pay the Open Access publication charges for this article were waivered by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

1. Cormack,B.P. and Struhl,K. (1992) The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. *Cell*, **69**, 685–696.
2. Kassavetis,G.A., Braun,B.R., Nguyen,L.H. and Geiduschek,E.P. (1990) *S. cerevisiae* TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. *Cell*, **60**, 235–245.
3. Bartholomew,B., Kassavetis,G.A. and Geiduschek,E.P. (1991) Two components of *Saccharomyces cerevisiae* transcription factor IIIB (TFIIIB) are stereospecifically located upstream of a tRNA gene and interact with the second-largest subunit of TFIIIC. *Mol. Cell. Biol.*, **11**, 5181–5189.
4. Chaussivert,N., Conesa,C., Shaaban,S. and Sentenac,A. (1995) Complex interactions between yeast TFIIIB and TFIIIC. *J. Biol. Chem.*, **270**, 15353–15358.
5. Moir,R.D., Sethy-Coraci,I., Puglia,K., Librizzi,M.D. and Willis,I.M. (1997) A tetratricopeptide repeat mutation in yeast transcription factor IIIC131 (TFIIIC131) facilitates recruitment of TFIIB-related factor TFIIIB70. *Mol. Cell. Biol.*, **17**, 7119–7125.
6. Joazeiro,C.A., Kassavetis,G.A. and Geiduschek,E.P. (1994) Identical components of yeast transcription factor IIIB are required and sufficient for transcription of TATA box-containing and TATA-less genes. *Mol. Cell. Biol.*, **14**, 2798–2808.
7. Kassavetis,G.A., Joazeiro,C.A., Pisano,M., Geiduschek,E.P., Colbert,T., Hahn,S. and Blanco,J.A. (1992) The role of the TATA-binding protein in the assembly and function of the multisubunit yeast RNA polymerase III transcription factor, TFIIIB. *Cell*, **71**, 1055–1064.
8. Kassavetis,G.A., Nguyen,S.T., Kobayashi,R., Kumar,A., Geiduschek,E.P. and Pisano,M. (1995) Cloning, expression, and function of TFC5, the gene encoding the B" component of the *Saccharomyces cerevisiae* RNA polymerase III transcription factor TFIIIB. *Proc. Natl Acad. Sci. USA*, **92**, 9786–9790.
9. Joazeiro,C.A., Kassavetis,G.A. and Geiduschek,E.P. (1996) Alternative outcomes in assembly of promoter complexes: the roles of TBP and a flexible linker in placing TFIIIB on tRNA genes. *Genes Dev.*, **10**, 725–739.
10. Starr,D.B. and Hawley,D.K. (1991) TFIID binds in the minor groove of the TATA box. *Cell*, **67**, 1231–1240.

11. Kim,J.L., Nikolov,D.B. and Burley,S.K. (1993) Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*, **365**, 520–527.
12. Kim,Y., Geiger,J.H., Hahn,S. and Sigler,P.B. (1993) Crystal structure of a yeast TBP/TATA-box complex. *Nature*, **365**, 512–520.
13. Juo,Z.S., Chiu,T.K., Leiberman,P.M., Baikalov,I., Berk,A.J. and Dickerson,R.E. (1996) How proteins recognize the TATA box. *J. Mol. Biol*., **261**, 239–254.
14. Horikoshi,M., Bertuccioli,C., Takada,R., Wang,J., Yamamoto,T. and Roeder,R.G. (1992) Transcription factor TFIID induces DNA bending upon binding to the TATA element. *Proc. Natl Acad. Sci. USA*, **89**, 1060–1064.
15. Starr,D.B., Hoopes,B.C. and Hawley,D.K. (1995) DNA bending is an important component of site-specific recognition by the TATA binding protein. *J. Mol. Biol*., **250**, 434–446.
16. Bucher,P. (1990) Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol*., **212**, 563–578.
17. Yamamoto,T., Horikoshi,M., Wang,J., Hasegawa,S., Weil,P.A. and Roeder,R.G. (1992) A bipartite DNA binding domain composed of direct repeats in the TATA box binding factor TFIID. *Proc. Natl Acad. Sci. USA*, **89**, 2844–2848.
18. Reddy,P. and Hahn,S. (1991) Dominant negative mutations in yeast TFIID define a bipartite DNA-binding region. *Cell*, **65**, 349–357.
19. Wong,J.M. and Bateman,E. (1994) TBP-DNA interactions in the minor groove discriminate between A:T and T:A base pairs. *Nucleic Acids Res*., **22**, 1890–1896.
20. Strubin,M. and Struhl,K. (1992) Yeast and human TFIID with altered DNA-binding specificity for TATA elements. *Cell*, **68**, 721–730.
21. Whitehall,S.K., Kassavetis,G.A. and Geiduschek,E.P. (1995) The symmetry of the yeast U6 RNA gene's TATA box and the orientation of the TATA-binding protein in yeast TFIIIB. *Genes Dev*., **9**, 2974–2985.
22. Grove,A., Kassavetis,G.A., Johnson,T.E. and Geiduschek,E.P. (1999) The RNA polymerase III-recruiting factor TFIIIB induces a DNA bend between the TATA box and the transcriptional start site. *J. Mol. Biol*., **285**, 1429–1440.
23. Persinger,J., Sengupta,S.M. and Bartholomew,B. (1999) Spatial organization of the core region of yeast TFIIIB-DNA complexes. *Mol. Cell. Biol*., **19**, 5218–5234.
24. Juo,Z.S., Kassavetis,G.A., Wang,J., Geiduschek,E.P. and Sigler,P.B. (2003) Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature*, **422**, 534–539.
25. Colbert,T. and Hahn,S. (1992) A yeast TFIIB-related factor involved in RNA polymerase III transcription. *Genes Dev*., **6**, 1940–1949.
26. Hoopes,B.C., LeBlanc,J.F. and Hawley,D.K. (1992) Kinetic analysis of yeast TFIID-TATA box complex formation suggests a multi-step pathway. *J. Biol. Chem*., **267**, 11539–11547.
27. Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*., **25**, 4876–4882.
28. Cloutier,T.E., Librizzi,M.D., Mollah,A.K., Brenowitz,M. and Willis,I.M. (2001) Kinetic trapping of DNA by transcription factor IIIB. *Proc. Natl Acad. Sci. USA*, **98**, 9581–9586.
29. Grove,A., Galeone,A., Yu,E., Mayol,L. and Geiduschek,E.P. (1998) Affinity, stability and polarity of binding of the TATA binding protein governed by flexure at the TATA Box. *J. Mol. Biol*., **282**, 731–739.
30. Perez-Howard,G.M., Weil,P.A. and Beechem,J.M. (1995) Yeast TATA binding protein interaction with DNA: fluorescence determination of oligomeric state, equilibrium binding, on-rate, and dissociation kinetics. *Biochemistry*, **34**, 8005–8017.
31. Petri,V., Hsieh,M. and Brenowitz,M. (1995) Thermodynamic and kinetic characterization of the binding of the TATA binding protein to the adenovirus E4 promoter. *Biochemistry*, **34**, 9977–9984.
32. Parvin,J.D., McCormick,R.J., Sharp,P.A. and Fisher,D.E. (1995) Pre-bending of a promoter sequence enhances affinity for the TATA-binding factor. *Nature*, **373**, 724–727.
33. Parkhurst,K.M., Brenowitz,M. and Parkhurst,L.J. (1996) Simultaneous binding and bending of promoter DNA by the TATA binding protein: real time kinetic measurements. *Biochemistry*, **35**, 7459–7465.
34. Petri,V., Hsieh,M., Jamison,E. and Brenowitz,M. (1998) DNA sequence-specific recognition by the *Saccharomyces cerevisiae* 'TATA' binding protein: promoter-dependent differences in the thermodynamics and kinetics of binding. *Biochemistry*, **37**, 15842–15849.
35. Wolner,B.S. and Gralla,J.D. (2001) TATA-flanking sequences influence the rate and stability of TATA-binding protein and TFIIB binding. *J. Biol. Chem*., **276**, 6260–6266.
36. Coleman,R.A. and Pugh,B.F. (1997) Slow dimer dissociation of the TATA binding protein dictates the kinetics of DNA binding. *Proc. Natl Acad. Sci. USA*, **94**, 7221–7226.
37. Faiger,H., Ivanchenko,M., Cohen,I. and Haran,T.E. (2006) TBP flanking sequences: asymmetry of binding, long-range effects and consensus sequences. *Nucleic Acids Res*., **34**, 104–119.
38. Calladine,C.R. and Drew,H.R. (1986) Principles of sequence-dependent flexure of DNA. *J. Mol. Biol*., **192**, 907–918.
39. Goodsell,D.S., Kaczor-Grzeskowiak,M. and Dickerson,R.E. (1994) The crystal structure of C-C-A-T-T-A-A-T-G-G. Implications for bending of B-DNA at T-A steps. *J. Mol. Biol*., **239**, 79–96.
40. Patikoglou,G.A., Kim,J.L., Sun,L., Yang,S.H., Kodadek,T. and Burley,S.K. (1999) TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev*., **13**, 3217–3230.
41. Cox,J.M., Hayward,M.M., Sanchez,J.F., Gegnas,L.D., van der Zee,S., Dennis,J.H., Sigler,P.B. and Schepartz,A. (1997) Bidirectional binding of the TATA box binding protein to the TATA box. *Proc. Natl Acad. Sci. USA*, **94**, 13475–13480.
42. Suzuki,M. and Yagi,N. (1995) Stereochemical basis of DNA bending by transcription factors. *Nucleic Acids Res*., **23**, 2083–2091.
43. Nikolov,D.B., Chen,H., Halay,E.D., Usheva,A.A., Hisatake,K., Lee,D.K., Roeder,R.G. and Burley,S.K. (1995) Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature*, **377**, 119–128.
44. Tan,S., Hunziker,Y., Sargent,D.F. and Richmond,T.J. (1996) Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature*, **381**, 127–151.
45. Geiger,J.H., Hahn,S., Lee,S. and Sigler,P.B. (1996) Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science*, **272**, 830–836.
46. Wu,J., Parkhurst,K.M., Powell,R.M., Brenowitz,M. and Parkhurst,L.J. (2001) DNA bends in TATA-binding protein-TATA complexes in solution are DNA sequence-dependent. *J. Biol. Chem*., **276**, 14614–14622.
47. Qian,X., Strahs,D. and Schlick,T. (2001) Dynamic simulations of 13 TATA variants refine kinetic hypotheses of sequence/activity relationships. *J. Mol. Biol*., **308**, 681–703.
48. Strahs,D., Barash,D., Qian,X. and Schlick,T. (2003) Sequence-dependent solution structure and motions of 13 TATA/TBP (TATA-box binding protein) complexes. *Biopolymers*, **69**, 216–243.
49. Dickerson,R.E. (1998) Sequence-dependent B-DNA conformation in crystals and in protein complexes. In Sarma,R.H. and Sarma,M.H. (eds), *Structure, Motion, Interaction and Expression of Biological Macromolecules*. Adenine Press, Schenectady, NY, pp. 17–36.
50. Wang,Y. and Stumph,W.E. (1995) RNA polymerase II/III transcription specificity determined by TATA box orientation. *Proc. Natl Acad. Sci. USA*, **92**, 8606–8610.