The TATA-binding protein (TBP) initiates assembly of transcription preinitiation complexes on eukaryotic class II promoters, binding to and restructuring consensus and variant “TATA box” sequences. The sequence dependence of the DNA structure in TBP-TATA complexes has been investigated in solution using fluorescence resonance energy transfer. The mean 5′-dye-3′-dye distance varies significantly among oligomers bearing the adenovirus major late promoter sequence (AdMLP) and five single-site variants bound to Saccharomyces cerevisiae TBP, consistent with solution bend angles for AdMLP of 76° and for the variants ranging from 30° to 62°. These solution bends contrast sharply with the corresponding co-crystal structures, which show ~80° bends for all sequences. Transcription activities for these TATA sequences are strongly correlated with the solution bend angles but not with TBP-DNA binding affinities. Our results support a model in which transcription efficiency derives primarily from the sequence-dependent structure of the TBP-TATA binary complex. Specifically, the distance distribution for the average solution structure of the TBP-TATA complex may reflect the sequence-dependent probability for the complex to assume a conformation in which the TATA box DNA is severely bent. Upon assumption of this geometry, the binary complex becomes a target for binding and correctly orienting the other components of the preinitiation complex.

The TATA-binding protein (TBP) binds to eukaryotic class II promoters at specific sequences of DNA of the consensus sequence TATA(a/t)A(a/t)N, nucleating assembly of the protein required for transcription. Atomic resolution co-crystal structures of complexes of DNA bearing consensus strong promoter sequences bound to Saccharomyces cerevisiae (1), Arabidopsis thaliana (2), and human (3, 4) TBPs are extremely similar, characterized by a TBP-induced ~80° bend in the DNA helix. TBP also binds to numerous variant TATA sequences, many of which occur naturally in promoters (5, 6). For 21 such single-point mutants of the adenovirus major late promoter (AdMLP) TATA box sequence, in vitro transcription activity was found to range from <1% to 107% of that of the reference AdMLP TATA sequence (6).

The wide range of observed transcription activities suggested that TBP does not bind similarly to all TATA elements. Gel electrophoresis circular permutation analysis of TBP-DNA complexes shows that the electrophoretic mobility of the complexes is TATA sequence-dependent, with bend angles from 34° to 106° inferred from the gel mobility patterns (7). In contrast, the co-crystal structures of 11 TATA sequence variants of varying affinity bound to A. thaliana TBP are all very similar, with the DNA helix bent as in the strong promoters (3, 8).

The present study was undertaken to further explore the TATA box sequence dependence of TBP binding and DNA structure using native, full-length S. cerevisiae TBP together with the AdMLP TATA sequence and five single-base-pair variant sequences. End-to-end distance distributions for these duplexes, free and TBP-bound, were extracted from measurements of time-resolved fluorescence emission in conjunction with fluorescence resonance energy transfer (FRET). Bend angles for the DNA within each of the TBP-DNA complexes were determined using three models. The reference AdMLP and five variant TATA sequences bound to TBP have significantly different mean end-to-end distances in solution. These distances are consistent with DNA bend angles ranging from 29.9° to 61.8° for the variant sequences and 76.2° for the native AdMLP. The latter bend angle is in excellent accord with the bends observed in the co-crystal structures. A strong correlation is observed between the solution bend angles and the transcription activities (6). These findings are consistent with the structure of TBP-TATA complexes being a principal determinant of TATA-box-dependent transcription activity. A model is proposed that reconciles the sequence dependence of bend angles and transcription activities measured in solution with the DNA structures observed in the co-crystals.

**Experimental Procedures**

Protein, DNA, and Solution Conditions—Full-length S. cerevisiae TBP was prepared as described previously (9, 10). The double-labeled 14 base oligonucleotides, with 5′-TAMRA and 3′-fluorescein, were as described previously (11). The specific sequences are shown in Table I. The corresponding single-labeled oligonucleotides (denoted 14-merF) were identical except that each had 3′-fluorescein but no
where \( I_0 \) and \( I_{th} \) are the fluorescence emission intensity of the donor in the absence and presence of acceptor, respectively, and the inverse of the 0th donor lifetime \( 1/\tau_0 \) equals \( k_0 + k_{th} \) and \( (1/\tau_0) = k_0 \) and \( k_{th} \) are the respective rate constants for fluorescence and nonradiative decay, and \( k_t \) is the rate constant for energy transfer. \( R_0 \) is the Förster distance, for which the efficiency of transfer is 0.5, and \( R_t \) is the donor lifetime uniquely associated with a particular value of \( R_0 \), and remains constant as long as the acceptor absorption and donor emission spectra remain unchanged. Thus, \( P(R) \) may be extracted from measurements of the fluorescence lifetime decay of the donor in the presence and absence of acceptor.

Fluorescence lifetime measurements were made in the time domain using a LaserStrobe spectrophotometer (Photon Technology International, Inc., Lawrenceville, NJ) with PTr dye PL481 to generate pulsed 488-nm excitation light. A 520-nm interference filter (Oriel Corp., Stratford, CT) between the sample compartment and the detector isolated the fluorescein emission.

The true fluorescence lifetime was extracted from the total emission, which included the tethers, their distance apart is variable and the donor decay depends upon the probability distribution, \( P(R) \), of all possible such distances (13, 14) as follows,

\[
I_0(t) = \int P(R) e^{-\frac{|t - \tau_0|}{\tau_0}(R)^2} \, dR \tag{Eq. 1}
\]

The determination of \( s \) was too low to achieve saturation of the duplex. TBP was added to 50 nM T*A3dpx*F to final concentrations of 153 and 178 nM (lower), 306 and 333 nM (intermediate), and 457 and 660 nM (higher), resulting in six different mixtures of free and TBP-bound A3dpx*F. Measurements of the fluorescence lifetimes of these mixtures were fixed in Eq. 2 at the mean values for the three composite curves previously obtained. The decay curves for the mixtures were analyzed globally in groups of three for four parameters: \( K_0 \), \( \alpha_1 \), \( \tau_1 \), and \( \gamma_1 \). The decay rates were determined as described in the previous section, and the resulting decay rates were used in the minimization procedure. The resulting four T*MLdpx*F decays were weighted to reflect their respective fractional populations, determined using \( K_a \). The values for \( R_{\text{free}} \) and \( R_{\text{bound}} \) were fixed at the previously determined values and the optimal values for \( K_a \) and \( a_{\text{bound}} \) obtained. The mean values and standard deviations for \( R_{\text{free}} \) and \( R_{\text{bound}} \) were determined as for the free duplex. The C7 data were analyzed identically to AdMLP. The same analysis obtained for T5, G6, and T8 but with 3 \times 3 matrices.

The determinations of \( P(R) \) for the free A3dpx*F were likewise obtained from the corresponding 4 \times 4 matrix. Because all solutions of bound DNA contained a small amount (<1% for AdMLP) of free duplex, each of the 16 probability distributions for the bound DNA were subsequently fit to the sum of two \( P(R) \) values corresponding to bound and free duplex. The former two classes of oligomers were high performance liquid chromatography/polyacrylamide gel electrophoresis and polyacrylamide gel electrophoresis-purified, respectively. Studies were conducted in 10 mM Hepes-HCl (pH 7.4), 100 mM KCl, 2.5 mM MgCl\(_2\), 1 mM CaCl\(_2\), and 1 mM dithiothreitol.

**Theory, Instrumentation, Data Acquisition, and Analysis—**Extensive discussions of Förster resonance energy transfer and its application to these studies have been published (Refs. 11–13 and references therein). Very briefly, FRET is the process whereby excited state energy is transferred nonradiatively from a donor to an acceptor fluorophore. Because both dyes are attached to an oligomer by flexible tethers, their distance apart is variable and the donor decay depends upon the probability distribution, \( P(R) \), of all possible such distances (13, 14) as follows:

\[
I_0(t) = \int P(R) e^{-\frac{|t - \tau_0|}{\tau_0}(R)^2} \, dR \tag{Eq. 1}
\]

The double- and single-labeled probes and unlabeled complementary oligomers were synthesized by Sigma-Genosys (The Woodlands, TX). The former two classes of oligomers were high performance liquid chromatography/polyacrylamide gel electrophoresis and polyacrylamide gel electrophoresis-purified, respectively. Studies were conducted in 10 mM Hepes-HCl (pH 7.4), 100 mM KCl, 2.5 mM MgCl\(_2\), 1 mM CaCl\(_2\), and 1 mM dithiothreitol.
decays, \( I_{d}(t) \), from these three mixtures were used in an expansion of Eq. 1.

\[
I_{d}(t) = F_{\text{free}}\left[ a_1 \exp\left( -\frac{t}{\tau_1} \right) + a_2 \exp\left( -\frac{t}{\tau_2} \right) \right] + F_{\text{bound}}\int_{1.5}^{12} P(R)\left( a_3 \exp\left( -\frac{t}{\tau_3} \right) \right) dR \quad \text{(Eq. 4)}
\]

with the values of the fractions of free (\( F_{\text{free}} \)) and bound (\( F_{\text{bound}} \)) T*A3dpx*F fixed according to \( K_c \). The values of \( a_1, a_2, \) and \( \tau_1, \tau_2, \) in the second and third terms on the right-hand side of Eq. 4 reflect bound donor-only decay and were fixed at the mean values previously determined, \( P(R) \) was then obtained using a matrix approach, analyzing separately all nine combinations of the three \( I_{d}(t) \) values, and, in the first term on the right-hand side, the three sets of \( a_1, \tau_1 \) and \( a_2, \tau_2 \) values from the three composite curves corresponding to free T*A3dpx*F. The resulting nine values for \( R_{\text{bound}} \) and \( \sigma_{\text{bound}} \) for T*A3dpx*F were averaged, and the S.D. was determined.

To confirm that the value of \( R_{\text{a}} \) remained essentially constant for all cases studied, \( R_{\text{a}} \) was determined independently for AdMLP and the T6 variant for both the free and TBP-bound duplexes, using the solvent refractive index of 1.332. The overlap integrals were determined independently as described (15) using emission spectra for free and bound MLdpx*F and T6dpx*F and absorption spectra for free and bound MLdpx*F and T6dpx*F. Emission and absorption spectra were collected on a steady-state fluorimeter (Photon Technology International, Inc., model A-1010) and a Hewlett-Packard diode array spectrophotometer (model HP8452A), respectively.

To establish sufficiently dye mobility consistent with \( s^2 = 2.3 \), time-resolved decay curves were also measured for free and TBP-bound MLdpx*F, T*Mdpx*F, T6dpx*F, and T*6dpx. Semi-cone angles for the dyes in each of these eight conditions were determined as described (18). Anisotropy decay measurements were made using the LaserStrobe spectrophotometer with 488- and 500-nm excitation light for fluorescein and TAMRA, respectively, and with the corresponding emission isolated by a 520-nm interference and a 530-nm long pass filter.

Thermosaturation of TBP—Because a time of \( \sim 80 \) min was required for one set of five replicates fluorescence lifetime measurements, the thermosaturation of TBP both free and DNA-bound was investigated. For the latter, the complex was formed using 1:1 TBP:DNA, both at 2.5 M, thus sampling the entire population of TBP molecules, and with the corresponding emission isolated by a 520-nm interference and a 530-nm long pass filter.

The equilibrium constant for each binding, as a duplex, to TBP was determined from the steady-state association equilibrium constant for each binding, as a duplex, to TBP was determined from the steady-state association constant of \( K_{eq} \) of native \( S. \) cerevisiae TBP binding these duplexes, determined by steady-state FRET, varied over a range of \( \sim 75 \times \) (Table 1). Of the variant sequences, only T5 binds TBP more tightly than the native AdMLP.

TBP Stability—Control experiments were conducted to demonstrate that a constant concentration of TBP-DNA complex was maintained throughout the course of a set of fluorescence lifetime measurements. The half-times for inactivation of the free and DNA-bound \( S. \) cerevisiae TBP binding preparation used in these studies (9, 10) were determined to be \( \sim 1 \) and \( \sim 10 \) h, respectively, at 30 °C under the experimental conditions of these studies. The protein remains fully active in DNA binding even after 24 h at 0 °C. These and similar results obtained by other assays contrast sharply with a recent report of the loss of the “vast majority” of the DNA binding activity of \( S. \) cerevisiae TBP after 0.5 min of incubation at 30 °C and all binding activity after 45 min even at 0 °C (21).

The dual-labeled 14-mer duplexes (denoted T*Mdpx*F and T6dpx*F) had 6-carboxy linker to 5’-TATA and 3’-fluorescein identical to those used previously (11, 13). These oligonucleotides differ only by a single base pair within the core sequence. The equilibrium association constants (\( K_{eq} \)) of native \( S. \) cerevisiae TBP binding these duplexes, determined by steady-state FRET, varied over a range of \( \sim 75 \times \) (Table 1). Of the variant sequences, only T5 binds TBP more tightly than the native AdMLP.

TBP Stability—Control experiments were conducted to demonstrate that a constant concentration of TBP-DNA complex was maintained throughout the course of a set of fluorescence lifetime measurements. The half-times for inactivation of the free and DNA-bound \( S. \) cerevisiae TBP binding preparation used in these studies (9, 10) were determined to be \( \sim 1 \) and \( \sim 10 \) h, respectively, at 30 °C under the experimental conditions of these studies. The protein remains fully active in DNA binding even after 24 h at 0 °C. These and similar results obtained by other assays contrast sharply with a recent report of the loss of the “vast majority” of the DNA binding activity of \( S. \) cerevisiae TBP after 0.5 min of incubation at 30 °C and all binding activity after 45 min even at 0 °C (21).

Numerical simulations mimicking the experimental conditions and incorporating these rates of TBP inactivation demonstrated the effects of TBP inactivation to be entirely negligible for at least the 80-min time period over which fluorescence lifetime data were acquired. This result is consistent with the experimental observation that the first and last curves were nearly identical for a given set of five replicate fluorescence lifetime decays (Fig. 1), showing no detectable loss of the protein-DNA complex with time.

\[ \text{Table I} \]

| Sequence | \( K_{eq} \) \( \mu M^{-1} \) |
|----------|-----------------|
| T*Mdpx*F | 168 (147, 191) |
| T5dpx*F  | 230 (196, 270) |
| G6dpx*F  | 76 (68, 88)    |
| C7dpx*F  | 60 (51, 70)    |
| T6dpx*F  | 29 (23, 33)    |
| A8dpx*F  | 3.1 (2.7, 3.5) |

\[ ^{2} \text{R. M. Powell, manuscript in preparation.} \]
\[ ^{3} \text{E. Jamison and M. Brenowitz, unpublished observations.} \]
\[ ^{4} \text{M. Daugherty and M. Fried, unpublished observations.} \]
\[ ^{5} \text{J. Wu, unpublished data.} \]
Variable DNA Bending in Solution TBP-TATA Complexes

End-to-End Distance Distributions, P(R), in Solution for Free and TBP-bound TATA Duplexes—The mean end-to-end distances (R̅) for the six duplexes, both free and TBP-bound, together with the corresponding values of σ for the distribution of distances, are listed in Table II. The values of R̅ were similar for all of the free duplexes except T*A3*MLdpx*F, differing by no more than 0.4 Å with an average value of 54.5 Å. The R̅ value of 53.4 ± 0.1 Å for the A3 variant is consistent with inherent curvature and/or flexibility related to the length of an uninterrupted A tract (23). A detailed study of these distance distributions for unbound DNA duplexes and their relationship to DNA sequence is in progress and will be reported elsewhere.

In contrast to the results obtained for the free DNA, the mean end-to-end distances for the TBP-bound duplexes varied significantly (Table II). The largest decrease in R̅ upon TBP binding, 7.3 Å, was measured for T*MLdpx*F. A decrease in the end-to-end distance was also observed upon TBP binding each of the variant duplexes, ranging from 4.9 to 0.1 Å.

The values of σ, the S.D. of the distribution, increased upon binding for all sequences, although not uniformly (Table II). A 5% increase in the breadth of the distribution for T*MLdpx*F contrasts with a 144% increase for T*T6dpx*F. The range of values of σ for the bound duplexes, 7.5 ± 0.1 to 10.5 ± 0.3 Å, greatly exceeds the confidence limits of the measurements.

Control experiments were conducted to ensure that the changes in the value of R̅ did not derive from changes in R0 due to sequence or to TBP binding. Semic cone angles, Ξ0, were determined for both fluorescein and TAMRA for the AdMLP and T6 duplexes, free and TBP-bound. (The transition moment of the fluorophore wobbles within a cone with the vertex at the center of the transition moment. The angle Ξ0 is half of the apical angle of this cone.) For these eight conditions, the semic cone angles ranged from 56° to 70° with an average value of 64°. Because the error for each angle was estimated to be ±7°, none of these angles differed significantly from the mean. The fast and slow rotational correlation times, τcorrespond to the free dye and to the macromolecule to which the dye is attached, respectively (18). For these eight conditions, τfast = 0.15 ± 0.03 ns (free and bound), τslow/free = 5 ± 2 ns, and τslow/bound = 23 ± 2 ns. These values of Ξ0 and τfast reflect a high degree of rotational freedom for both dyes, free and bound, for two sequences with disparate values for R̅. Furthermore, the fluorophores are in very similar environments for the six duplexes, because all sequences are identical for the same material (dashed line). The latter curve was obtained after ~60 min and shows no trend toward the time progression of the free duplex, due both to the stability of the bound protein and to the large excess of TBP.

The independently collected 3'-fluorescein emission and 5'-TAMRA absorption spectra for the free AdMLP and T6 duplexes were invariant, yielding identical overlap integrals and R0 = 61.0 Å. The integrals were likewise invariant for TBP-bound AdMLP and T6, with R0 = 61.2 Å, an increase of 0.3% upon binding. These values of R0free and R0bound were therefore assumed for the other four sequences. The results of these control experiments confirm that the sequence dependence of R̅ does not derive from changes in R0.

TATA Sequence-dependent Solution Bend Angles for TBP-bound Duplexes—We have shown previously that the P(R) determined for TBP-bound T*MLdpx*F using FRET fluorometry is generally consistent with the bent DNA observed in the co-crystals bearing strong promoter sequences (9). The relationship of R̅ to the solution bend angle has been further explored by consideration of three models of DNA bending (Fig. 2). Although these models do not account for phasing of the dyes and unwinding of the duplex, the present data do not allow the critical testing of more complex models incorporating such detailed parameters of DNA structure.

For these three models, the smooth bend and single central bend models (Fig. 2, C and A, respectively) correspond to the upper and lower limits for the bend angles for a given ratio of

![Figure 1](http://www.jbc.org)

**Figure 1.** Typical time-resolved fluorescence decays, for T*MLdpx*F free in solution (dotted line) and TBP-bound. The decay collected initially for bound T*MLdpx*F (solid line), following equilibration at 30 °C of 50 nm duplex with 440 nm TBP, is very similar to the fifth decay collected in the data set on the same material (dashed line). The latter curve was obtained after ~60 min and shows no trend toward the time progression of the free duplex, due both to the stability of the bound protein and to the large excess of TBP.

![Table II](http://www.jbc.org)

**Table II**

| Sequence | Free duplex, P(R)free | TBP-DNA complex, P(R)complex |
|----------|----------------------|-----------------------------|
| AdMLP    | 54.4 ± 0.1 8.1 ± 0.1 | 47.1 ± 0.1 8.5 ± 0.1 |
| T5       | 54.5 ± 0.1 4.6 ± 0.2 | 49.6 ± 0.1 7.5 ± 0.1 |
| G6       | 54.6 ± 0.2 4.7 ± 0.2 | 50.6 ± 0.1 9.8 ± 0.1 |
| C7       | 54.7 ± 0.1 6.8 ± 0.2 | 51.2 ± 0.1 10.4 ± 0.1 |
| T6       | 54.3 ± 0.1 4.3 ± 0.2 | 52.5 ± 0.1 10.5 ± 0.3 |
| A3       | 53.4 ± 0.1 7.4 ± 0.3 | 53.3 ± 0.2 10.3 ± 0.4 |
TABLE III
Derived solution bend angles according to three bending models for the six duplex DNAs bound to TBP

|         | Single central bend | Two-kink bend | Smooth bend |
|---------|---------------------|---------------|-------------|
| AdMLP   | 60.1                | 76.2 ± 0.2    | 105.0       |
| T5      | 49.0                | 61.8 ± 0.2    | 85.3        |
| G6      | 44.1                | 56.5 ± 0.3    | 76.8        |
| C7      | 41.2                | 52.3 ± 0.3    | 71.7        |
| T6      | 29.6                | 36.9 ± 0.2    | 51.4        |
| A3      | 24.1                | 29.9 ± 7.2    | 41.8        |

$R_{\text{bound}}/R_{\text{free}}$ (Table III). Both models are universal in that the calculated bend angle depends only on the ratio of $R_{\text{bound}}/R_{\text{free}}$ for bent versus linear DNA and is independent of DNA length.

Because of the nature of the DNA bend observed within the TBP-TATA complex, with sharp kinks at either end of the TATA sequence (1, 2), a two-kink model (Fig. 2B) has also been considered. The bend angles derived from this model are intermediate between the smooth and single central bend models, and depend on both the total length of the oligonucleotide and the position of the kinks. The bend angles that derive from the ratio of $R_{\text{bound}}/R_{\text{free}}$ corresponding to each model are as follows,

**Smooth bend model:**

\[
\frac{R_{\text{bound}}}{R_{\text{free}}} = \frac{2 \sin(\alpha)}{\alpha}
\]  
(Eq. 5)

**Two-kink model:**

\[
\frac{R_{\text{bound}}}{R_{\text{free}}} = \frac{(R_{\text{free}} + L_2 \cos(\alpha/2) + L_1}{R_{\text{free}}}
\]  
(Eq. 6)

**Single central bend model:**

\[
\frac{R_{\text{bound}}}{R_{\text{free}}} = \cos(\alpha/2)
\]  
(Eq. 7)

where $\alpha$ and $L_2$ are as described in Fig. 2 and all distances are in angstroms. For all models, the linear distance is assumed to be the measured $R$ for the free duplex.

The DNA bend observed in all the co-crystal structures is relatively smooth, with most of the total bend occurring at the flanking Phe intercalation sites and the remainder in-between. For TBP-bound T*MLpx*F, the solution bend angle associated with the two-kink model, $76.2 \pm 0.2^\circ$, closely corresponds to that observed in the co-crystals (Table III). The solution angle derived from the smooth bend model, $105.0^\circ$, is significantly larger, and from the single central bend model, $60.1^\circ$, significantly smaller. As we show in the accompanying paper (24), the close correspondence of the AdMLP bend in solution and in the co-crystal demonstrates clearly the appropriateness of the two-kink model to the interpretation of the FRET data.

The values of $R$ for the variant sequences are consistent with solution bend angles, determined using the two-kink model, ranging from $29.9^\circ$ (for A3) to $61.8^\circ$ (for T5). With the exception of the A3 variant, the solution bend angles reported in Table III derive from ratios of $R_{\text{bound}}/R_{\text{free}}$ with the latter specific for the given sequence. Because free T*A3px*F appears to have inherent curvature, the bend angle reported for TBP-bound A3 was determined using the average value of $R_{\text{free}}$ for the other five sequences of 54.5 Å. It is notable that the A3 sequence, alone among those studied, is not known to occur naturally and has transcription activity that is <1% of the AdMLP (6). The differences in bend angles observed among the AdMLP and variant promoter sequences are highly significant due to the high precision in the values of $R$ (see “Experimental Procedures”).

The similarity of the AdMLP bend angles determined in the co-crystal and in solution using the two-kink model demonstrates the adequacy of this model in describing the overall AdMLP conformation, without detailed consideration of DNA structure such as helical unwinding. Comparisons of the variant sequences to the reference AdMLP thus focused on this model, with the larger values of $R_{\text{bound}}$ for the variants interpreted as relatively smaller bends. These larger $R$ values could in principle also arise from an increase in the contour length of the DNA due to increased variant helical unwinding. However, maximum unwinding would be expected to be associated with

*Using instead the average value of $R_{\text{free}}$, 54.5 Å, as the reference yields bend angles that are all within 2.5° of those reported.
maximum bending to relieve the strain introduced by bending. Furthermore, the inter-phenylalanine distances (measured from the a carbons) differ by <1 Å in the TBP crystal structures with and without bound DNA bent to 80° (1, 2), the conformations corresponding to the maximum structural distortion. The distance between the kink sites, L* was therefore held constant in Eq. 6 for all sequences. Differences in R* are then assumed to derive primarily from differences in the details of phenylalanine intercalation in the minor groove at the kink sites, with greater penetration resulting in increased compression of the major groove and greater bending.

Correlation between Bend Angle and the Breadth of the Distribution—Because the end bases and linker arms are identical in all oligonucleotides studied, the differences in the values of s are assumed to derive primarily from the duplex DNA rather than the linker arms. The changes in s, Δs, upon TBP binding, rather than s itself, provide the most informative comparison among the sequences. The relationship between Δs and derived bend angle is shown in Fig. 3. The native AdMLP sequence (T*MLdpx*F), with the largest bend angle, has the smallest increase in s of only 0.4 Å upon TBP binding.

**DISCUSSION**

Time-resolved fluorescence resonance energy transfer provides a rigorous approach to the determination of the structure and dynamics of macromolecules in solution. The primary experimental findings from this work are 1) the existence in solution of DNA sequence-dependent differences in the trajectory of the DNA as it passes through TBP-TATA complexes and 2) the inverse correlation between the observed DNA bend angle and the breadth of the corresponding distance distribution.

**DNA Bend Angles in TBP-TATA Complexes and the Corresponding Probability Distributions Are DNA Sequence-dependent**—The FRET data clearly demonstrate sequence-dependent differences in the trajectory of the DNA as it passes through TBP-DNA complexes. In sharp contrast to this result and similar conclusions drawn from circular permutation and DNA phasing studies (7, 25), eleven variant TATA sequences bound to TBP, including all of the sequences in this study, have essentially identical ~80° DNA bends in the atomic resolution structures determined for TBP-DNA co-crystals (3, 8). These contrasting results are accommodated within a two-state allosteric model, based on an equilibrium between transcriptionally active and inactive TBP-DNA conformations (discussed below). The apparent conundrum presented by the solution and co-crystal structures is then definitively explained in the accompanying paper (24).

Also important for consideration of the underlying mechanism of this observation are the differences in the breadths of the corresponding distance distributions provided by the time-resolved FRET data. Clearly both structure and dynamics contribute to TBP-TATA function. The AdMLP sequence alone shows only a slight increase in the value of s, the S.D. of the end-to-end distance distribution, upon TBP binding. A plausible hypothesis is that the complementarity of the protein-duplex interface confines the helix and restricts additional motion. This slight increase in the breadth of the distribution for the tightly bound AdMLP may derive from the presence of multiple conformers at equilibrium, each with bent DNA but differing, for example, in the extent of phenylalanine intercalation (11). An integrated hydroxyl radical footprinting and molecular dynamics study of the TBP-AdMLP interface supports this view of its dynamic nature (26).

The variant sequences show a general trend toward increasingly broader distributions as the extent of bending decreases, up to Δs = 6.2 Å for the T6 variant. The inverse correlation between bending extent and distribution broadening may derive from the increasing misfit along the protein-DNA interface as helical bending decreases, including retention of solvent molecules at the interface. Indeed, complexes of TBP with the variant duplexes may be present in multiple conformations with the DNA bent very differently among those conformers, as discussed further in the following section. The broadened distribution would then result from equilibrium exchange among such conformers occurring on a time scale that is slow relative to the nanosecond time scale of the measurements, i.e. microseconds. In this case, the broader distribution of distances would not derive from any high frequency torsional and bending motions of the duplex that occur on time scales faster than nanoseconds, because such motion would be averaged out in these measurements (27).

**A Bi-modal Distribution Model Reconciling the Solution and Co-crystal Bend Angles**—A two-state model is hypothesized, unifying into a coherent perspective the sequence-dependent solution bend angles reported herein and the x-ray results in which only an AdMLP-like structure was crystallized. Each variant duplex bound to TBP is proposed to exist in two conformations, one (conformer_ML) with the DNA bound and bent as in the AdMLP-TBP complex and the other (conformer_T1) with the DNA significantly less bent (Fig. 4A). Only conformer_T1 has the correct geometry to allow binding of subsequent transcription proteins and effect measurable mRNA synthesis. Conformer_ML is transcriptionally inactive and has the same
Korean distributions corresponding to conformer ML and simple relationship among the variants to explain their conditions for the crystallizations differed from those employed for the crystallized (except A3 (8)), although the solution equilibrium for conformer TI is fast relative to subsequent binding processes, the transcription factors “see” and appear macroscopically to bind to an average TBP-DNA structure that is sequence-dependent. The model predicts that the more AdMLP-like the average binary structure, the more efficiently transcription will proceed. Implicit in this model is a correspondence between the structure of the TBP-TATA complex and transcription activity, which is explored further below.

**Minimal Correspondence of TBP-DNA Complex Lifetime to Bend Angle or Transcriptional Activity—Hawley and coworkers (7), inferring bend angles from gel mobility shifts for TBP-bound AdMLP and eight variant sequences, also observed sequence-dependent differences in bend angles. However, for the sequences common to both studies, those angles differ from those reported herein in magnitude, by up to a factor of two, but more significantly, in the ordering of sequences by decreasing bend. Although a correlation was asserted between bend angles inferred from circular permutation analysis and TBP-TATA complex stability (7), careful inspection of those data reveal a minimal correspondence between these two properties. A plot of the lifetime of the TBP-DNA complex versus bend angle from Table I (7) shows no general linear correlation (correlation = 0.76, coefficient of determination = 0.59); rather, the data form two distinct sets. The first of these sets of five sequences is composed of unstable TBP-TATA complexes, with lifetimes ≤ 0.08 that of the wild type, but with bend angles ranging from <34° to 80°. The second set of five sequences, constituting a step function relative to the first set, includes only severe bends, from 80° to 106°, but lifetimes that vary 23-fold, from 0.08 to 1.85 that of the wild type. This conclusion is further supported by the recent work of Bareket-Samish et al. (25), who report no correlation between TBP-TATA complex stability and DNA bend angles determined similarly using gel phasing analysis.

**DNA Bends in TBP-DNA Complexes Are Highly Correlated with Relative Transcription Activity**—The correlation between the solution bend angles determined in the present study and the corresponding in vivo and in vitro transcription activities...
reported by Wobbe and Struhl (6) are shown in Figs. 5, A and B, respectively. The same correlation is observed upon comparison with either the HeLa TFIID or yeast TBP in in vitro transcription assays. Two possible explanations for this correlation present themselves. First, the observed differences in transcription activity are structurally based, resulting fundamentally from the sequence-dependent differences in the DNA bend angles in the binary complexes, or second, they derive mentally from the sequence-dependent differences in the DNA binding affinity. Then, for example, were the AdMLP sequence 95% ally saturated in accord with their respective binding constants. The concentrations of HeLa and yeast TBP used in the in vitro assays were reported to be saturating under the experimental conditions of those studies (6). We therefore conclude that the >100-fold differences observed in transcription efficiency could not have arisen from differences in TBP-DNA affinity.

Suppose, however, that only the tightly bound AdMLP sequence was saturated and the variant sequences were fractionally saturated in accord with their respective binding constants, so that transcription activity did reflect differences in affinity. Then, for example, were the AdMLP sequence 95% bound (as a lower limit), the transcription activity for the T6 sequence would be 86% that of AdMLP, based on the \( K_d \) values shown in Table I. In contrast, the experimentally observed transcription activity for T6 was only 10% that of AdMLP (6). Thus, several independent lines of evidence support the conclusion that differences in TBP-DNA binding affinity cannot account for the observed differences in transcription efficiency.

In contrast, a significant correlation is observed between the solution bend angles and transcription activity. Wobbe and Struhl (6) similarly concluded that the in vitro activity of a TATA element is directly affected by the binary TBP-TATA structure. This conclusion was based on the close similarity between the in vitro activity of yeast TBP (and human TFIID) and transcription activity in yeast cells. The strong correspondence between the solution geometry of the TBP-DNA complex and transcription activity is further supported by a comparison of Figs. 4B and 5B. The relationship between transcription efficiency and bend angle is strikingly similar to the relationship between the fractional population of the allosteric conformer, and bend angle. The extent to which conformer \( \text{ML} \) is populated, for a given sequence, thus closely corresponds to the relative transcription activity.

The relatively large values of \( \sigma \) determined herein for the bound duplexes with less favorable TATA box sequences are consistent with low frequency DNA flexibility within the binary complexes. Such duplex motions cannot be effectively distinguished from multiple conformations (29). The observed correlation between the extent of DNA bending and transcription activity thus leads us to propose that the probability for a given TBP-TATA complex to assume the conformation required for binding of subsequent proteins determines the corresponding transcription efficiency. For the bound variants, as the deviation from –80° increases, severe distortions of the duplex DNA to approach 80° become increasingly less probable. In terms of such fluctuations, a dependence of transcription efficiency on the average conformation of the binary TBP-promoter complex seems reasonable. Both biochemical and crystallographic results show that flanking sequences up- and downstream of the TATA box are contacted by TFIIB (30, 31) and TFIIB (32–34), with TFIIB contacting both. Appropriately bent DNA in the TBP-DNA target may thus be critical for formation of stable ternary and quaternary complexes involving these proteins.

The trajectories of the helical axes resulting from different bends diverge rapidly (Fig. 6). For example, for a 14-bp duplex centered on the TATA box, the difference in the 5′-3′ distance between a 40° and an 80° bend is 4 Å. Extension of the duplex by only 6 bp up- and downstream, for example, more than triples that difference, from 4 to 13 Å. TBP-bound T6 and AdMLP have angles of –40° and –80°, respectively, and 6-bp extensions correspond generally to the flanking contact regions for TFIIB and -B. Formation of a stable higher-order structure is thus predicted to be less probable for the TBP-T6 complex than for the TBP-AdMLP complex, due to the spatial requirements.

In drawing a correlation between the apparent bend angle and transcriptional activity, however, consideration must be given to the experimental conditions of the respective studies. The in vitro transcription assays were performed in the presence of osmolyte (6, 35). As shown in the accompanying paper (24), the conformations of some bound variant sequences are
sensitive to the presence of osmolyte. Because a significant correlation is observed between bend angle and transcriptional activity both in vitro and in vivo (Fig. 5, A and B), it is plausible that the extremely small differences in energy between conformers for these sequences (24) are compensated in osmolyte by protein-protein interactions among multiple transcription factors. How the binding of even one additional transcription protein, in osmolyte, might affect the equilibrium among sequence-dependent TBP-DNA conformers is not known. Thus, effects of osmolytes on the conformation of the binary complex within multiprotein complexes require further exploration.

However, unequivocal new insight is provided by elucidation of the solution structures of TBP-AdMLP and TBP-A3. These binary complexes with the high and low extremes of the observed bend angles correspond to the high and low extremes of transcriptional activity. The solution geometries of these two complexes are insensitive to the presence of osmolyte (24) and establish clearly the relationship between transcription activity and the structure of the binary complex.

Conclusions—The geometries of the TBP-bound variant TATA sequences in solution vary significantly and differ from their corresponding co-crystal structures. These solution conformations are consistent with DNA bend angles ranging from ~30 to ~76° based on a two-kink bending model. A strong correlation between the solution bend angles and relative transcription activity, but not with TBP-DNA affinity, is observed. This correlation is particularly notable, because efficient transcription requires complex geometric relationships among many proteins and to summarize such complexity with a single, simple bend angle must be, to some extent, an oversimplification.

This model contrasts with models in which the TBP-DNA binary complex structure is conserved (8) and sequence-dependent differences in transcription efficiency derive primarily from sequence-dependent differences in the stability of that complex (7, 8). Our results support a model in which transcription efficiency derives in significant part from the sequence-dependent structure of the TBP-TATA binary complex. More specifically, the distance distribution for the average solution structure of the TBP-TATA complex may reflect the sequence-dependent probability for the complex to assume a conformation in which the TATA box DNA is severely bent. Upon assumption of this geometry, the binary complex becomes a target for binding and correctly orienting the other components of the preinitiation complex.

Acknowledgment—We thank Stephen Burley for communication of results prior to publication.

REFERENCES

1. Kim, Y., Geiger, J. H., Hahn, S., and Sigler, P. B. (1993) *Nature* **365**, 512–519
2. Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993) *Nature* **365**, 520–527
3. Nikolov, D. B., Chen, H., Halay, E. D., Hoffmann, A. R., Roeder, R. G., and Burley, S. B. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 4862–4867
4. Juo, Z. S., Chiu, T. K., Leiberman, P. M., Babilin, J. B., Berk, A. J., and Dickerson, R. E. (1996) *J. Mol. Biol.* **261**, 239–254
5. Hahn, S., Buratowski, S., Sharp, P., and Gurvance, L. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 7429–7435
6. Wobbe, C. R., and Struhl, K. (1990) *Mol. Cell. Biol.* **10**, 3859–3867
7. Starr, D. B., Hoopes, B. C., and Hawley, D. K. (1995) *J. Mol. Biol.* **250**, 434–446
8. Patitzion, G. A., Kim, J. L., Sun, L., Yang, S. H., Kodadek, T., and Burley, S. K. (1999) *Genes Dev.* **13**, 3217–3230
9. Parkhurst, K. M., Brenowitz, M., and Parkhurst, L. J. (1996) *Biochemistry* **35**, 7429–7435
10. Petri, V., Hsieh, M., and Brenowitz, M. (1995) *Biochemistry* **34**, 9977–9984
11. Parkhurst, K. M., Richards, R. M., Brenowitz, M., and Parkhurst, L. J. (1999) *J. Mol. Biol.* **289**, 1327–1341
12. Parkhurst, K. M., and Parkhurst, L. J. (1995) *Biochemistry* **34**, 285–292
13. Parkhurst, K. M., and Parkhurst, L. J. (1995) *Biochemistry* **34**, 293–300
14. Cantor, C. R., and Pechuakas, P. (1971) *Proc. Natl. Acad. Sci. U. S. A.* **68**, 2099–2101
15. Durbin, J., and Watson, G. S. (1950) *Biometrika* **37**, 409–428
16. Durbin, J., and Watson, G. S. (1951) *Biometrika* **38**, 159–178
17. Hamburg, M. (1974) *Basic Statistics: A Modern Approach*, pp. 285–287, Harcourt Brace Jovanovich, Inc., New York
18. Bucci, E., and Steinier, R. F. (1988) *Biophys. Chem.* **30**, 199–224
19. Daugherty, M. A., Brenowitz, M., and Fried, M. G. (1999) *J. Mol. Biol.* **285**, 1389–1399
20. Daugherty, M. A., Brenowitz, M., and Fried, M. G. (2000) *Biochemistry* **39**, 4869–4880
21. Jackson-Fisher, A. J., Burma, S., Portnoy, M., Schneeweis, L. A., Coleman, R. A., Mitra, M., Chitikila, C., and Pugh, B. F. (1999) *Biochemistry* **38**, 11340–11348
22. Parkhurst, L. J., and Parkhurst, K. M. (1994) *Proc. Soc. Photo-Opt. Instrum. Eng.* **2137**, 475–483
23. Kuo, H.-S., Wu, H.-M., and Crothers, D. M. (1986) *Nature* **320**, 501–506
24. Wu, J., Parkhurst, K. M., Powell, R. M., and Parkhurst, L. J. (2001) *J. Biol. Chem.* **276**, 14623–14627
25. Bareket-Samish, A., Cohen, I., and Haran, T. E. (2000) *J. Mol. Biol.* **299**, 965–977
26. Pastor, N., Weinstein, H., Jamison, E., and Brenowitz, M. (2000) *J. Mol. Biol.* **304**, 55–68
27. Okonogi, T. M., Reese, A. W., Alley, S. C., Hopkins, P. B., and Robinson, B. H. (1999) *Biochem. J.* **326**, 535–542
28. Yang, M., and Millar, D. P. (1996) *Biochemistry* **35**, 7959–7967
29. Naimushin, A. N., Fujimoto, B. S., and Schurr, J. M. (2000) *Biochemistry* **39**, 250, 235–250
30. Lagrange, T., Kim, T.-K., Orphanides, G., Ebright, Y. W., Ebright, R. H., and Reinberg, D. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10620–10625
31. Coulombe, B., Li, J., and Greenblatt, J. (1994) *J. Biol. Chem.* **269**, 26, 19962–19967
32. Lee, S., and Hahn, S. (1995) *Nature* **376**, 609–612
33. Lagrange, T., Kapanidis, A. N., Tang, H., Reinberg, D., and Ebright, R. H. (1998) *Genes Dev.* **12**, 34–44
34. Tsai, F. T. F., and Sigler, P. B. (2000) *EMBO J.* **19**, 25–36
35. Sawadogo, M., and Roeder, R. G. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 4384–4389
DNA Bends in TATA-binding Protein·TATA Complexes in Solution Are DNA Sequence-dependent
Jiong Wu, Kay M. Parkhurst, Robyn M. Powell, Michael Brenowitz and Lawrence J. Parkhurst

J. Biol. Chem. 2001, 276:14614-14622.
doi: 10.1074/jbc.M004402200 originally published online January 26, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M004402200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 34 references, 11 of which can be accessed free at
http://www.jbc.org/content/276/18/14614.full.html#ref-list-1