6-25-1999

Intermediate species possessing bent DNA are present along the pathway to formation of a final TBP-TATA complex

Kay M. Parkhurst
University of Nebraska - Lincoln

Robyn M. Richards
University of Nebraska - Lincoln

Michael Brenowitz
Department of Biochemistry, Albert Einstein College of Medicine, Bronx NY

Lawrence J. Parkhurst
University of Nebraska - Lincoln, lparkhurst1@unl.edu

Published in *Journal of Molecular Biology* 289:5 (June 25, 1999), pp. 1327–1341. doi:10.1006/jmbi.1999.2835

Submitted October 5, 1998; revised April 14, 1999; accepted April 22, 1999.

Intermediate species possessing bent DNA are present along the pathway to formation of a final TBP-TATA complex

Kay M. Parkhurst[1], Robyn M. Richards[1], Michael Brenowitz[2], and Lawrence J. Parkhurst[1] *

[1] *Department of Chemistry University of Nebraska–Lincoln, Lincoln NE, 68588-0304*

[2] *Department of Biochemistry Albert Einstein College of Medicine, Bronx NY, 10461*

* Corresponding author. email: lparkhurst@unl.edu

Abstract: Binding of the TATA-binding protein (TBP) to the "TATA" sequences present in the promoters of eukaryotic class II genes is the first step in the sequential assembly of transcription pre-initiation complexes. Myriad structural changes, including severe bending of the DNA, accompany TBP-TATA complex formation. A detailed kinetic study has been conducted to elucidate the mechanistic details of TBP binding and DNA bending. The binding of *Saccharomyces cerevisiae* TBP to the adenovirus major late promoter (AdMLP) was followed in real-time through a range of temperatures and TBP concentrations using fluorescence resonance energy transfer (FRET) and stopped-flow mixing. The results of association and relaxation kinetics and equilibrium binding experiments were analyzed globally to obtain the complete kinetic and energetic profile of the reaction. This analysis reveals a complex mechanism with two intermediate species, with the DNA in the intermediates apparently bent similarly to the DNA in the final complex. TBP binding and DNA bending occur simultaneously through the multiple steps of the reaction. The first and third steps in this sequential process show nearly identical large increases in both enthalpy and entropy, whereas the middle step is highly exothermic and proceeds with a large decrease in entropy. The first intermediate is significantly populated at equilibrium and resembles the final complex both structurally and energetically. It is postulated that both this intermediate and the final complex bind transcription factor IIB in the second step of pol II pre-initiation complex assembly. A consequence of such a reactive intermediate is that the rate of assembly of transcriptionally competent pre-initiation complexes from bi-directionally bound TBP is greatly increased.

Keywords: TBP, TATA, DNA-protein interaction, fluorescence resonance energy transfer (FRET), transcription

Abbreviations: TBP, TATA-binding protein; FRET, fluorescence resonance energy transfer; TAMRA, carboxytetramethylrhodamine; T*ML_{dpx}*F, double-labeled duplex containing the adenovirus major late sequence (TAMRA-$^{5'}$GGGC**TATAAAAG**GG$^{3'}$-fluorescein); ML_{dpx}, unlabeled duplex; PIC, pre-initiation complex; AdMLP, adenovirus major late promoter

Introduction

The first step in the initiation of gene transcription in eukaryotes by RNA polymerase II is the binding of the TATA-binding protein (TBP) to specific sequences of promoter DNA of the consensus sequence TATAa/tAa/tN ("TATA boxes"; [Burley and Roeder 1996, Hernandez 1993, Maldonaldo and Reinberg 1995, McKnight 1996 and Horikoshi et al 1992]). Studies of TBP from several species have been conducted. Its binding to DNA containing TATA box sequences results in dramatic structural changes, particularly in the DNA, including minor groove widening, major groove compression and unwinding of the DNA helix by ~120°. Two pairs of phenylalanine residues intercalate at each end of the TATA box, resulting in two "kinks" in the DNA helix and an overall bend in the helical axis to ~108° ([Juo et al 1996, Kim and Burley 1994, Kim

et al 1993, Kim et al 1993, Nikolov et al 1996, Parkhurst et al 1996 and Starr et al 1995]). Although DNA distortions induced by the intercalation of amino acid residues have been observed in several DNA-protein complexes [Werner et al 1996], the structure of the TBP-TATA complex is unique in several aspects.

Such profound changes in macromolecular structure upon the binding of TBP to a TATA sequence suggest a complex reaction mechanism for this process. The fact that TBP-TATA association reactions proceed at rates significantly slower than diffusion limited rates, in contrast to most DNA-binding proteins that have been studied, is well documented (e.g. [Hoopes et al 1992, Perez-Howard et al 1995, Petri et al 1995, Petri et al 1998, Parkhurst et al 1996 and Coleman and Pugh 1997]). Recent work in our laboratory, based on Förster resonance energy transfer (FRET) and fluorescence stopped-flow methodology, showed that for the adenovirus major late promoter, *Saccharomyces cerevisiae* TBP binding and DNA bending occur simultaneously [Parkhurst et al 1996]. This observation rules out a two-step mechanism in which the initial binding of DNA is followed by slow bending. In these initial FRET kinetics studies conducted at 30°C, subtle deviations from single exponential behavior were noted (see Discussion and Figure 5, inset) that suggested intermediate species might be discernible if the association reaction were followed as a function of temperature and TBP concentration. In addition, two TATA box sequences tightly bound by TBP, the adenovirus major late (TATAAAAG) and E4 (TATATATA) promoters, exhibit different equilibrium and kinetic energetic profiles [Petri et al 1998]. Remarkably, the dissociation "rate constants" calculated from the association equilibrium and kinetic measurements in that comparative study display no activation energy (see Figure 4), a result confirmed by direct determinations in the present work. Taken together, these data suggest a complex reaction pathway for TBP-DNA interactions that is both DNA sequence and temperature dependent.

It was to determine the molecular origins of these characteristics of TBP-promoter interactions that a "real-time" FRET kinetics study has been conducted. The association kinetics of the *S. cerevisiae* TBP binding the major late promoter were obtained as a function of TBP concentration and temperature using stopped-flow FRET. These data were globally fit with data from relaxation (overall dissociation), kinetic and equilibrium binding determinations obtained using steady-state FRET over the same temperature range to reveal the complete kinetic and energetic profile of this process. A model with two intermediates, both with DNA apparently bent to essentially the same extent as the final TBP-DNA complex, best describes the binding of TBP to the TATA box of the major later promoter. Although TBP binding and DNA bending indeed occur simultaneously [Parkhurst et al 1996], individual steps in the reaction pathway are discernible, consistent with the extensive structural changes that accompany complex formation. One of the intermediates is the dominant species for much of the association reaction, and remains as a significant component of the final equilibrium population of TBP-DNA. The presence of an abundant and long-lived intermediate in the association reaction opens the possibility of kinetic control of subsequent steps in the assembly of pre-initiation transcription complexes.

Results

In order to clarify the presentation of the complex reaction kinetics that have been observed, the terminology used to describe such mechanisms is briefly reviewed. For a single-step first-order (or pseudo first-order) reaction that proceeds to completion, the time-course of the change in an observable property (e.g. fluorescence intensity) follows simple first-order decay. In this simple case, the "decay constant" corresponds to the microscopic rate constant. However, if the reaction is reversible and proceeds to equilibrium, the observed decay constant is not simply a "rate constant" but is the sum of the forward and reverse rate constants.

In contrast, for any multiple-step first-order (or pseudo first-order) reaction, the observed decay constants do not directly correspond to the microscopic rate constants. Instead, the observed constants are referred to as relaxation constants or eigenvalues (λ_i) that describe the time-dependent "relaxation" of a system toward equilibrium. (These constants are denoted eigenvalues, since they are the eigenvalues of the rate constant matrix for the multiple steps of the reaction.) The number of non-zero eigenvalues associated with a reaction is equal to the number of thermodynamically independent steps in the mechanism. In general, the concentration of each species and each observable property is a linear combination of exponentials ($\Sigma\ \alpha_i\ \exp{-\lambda_i t}$), where each λ_i is a function of all of the rate constants in the mechanism and α_i is the amplitude of the exponential *i*. Whether the *i* th eigenvalue can be detected in an observed decay curve depends upon whether it is distinguishable from the other eigenvalues and whether the associated amplitude is large enough to allow detection. From these complex decays it may be possible to extract the microscopic rate constants associated with a given model.

Fluorescence stopped-flow measurements

A total of 45 stopped-flow kinetic curves were collected for the binding of TBP to the double-labeled DNA duplex (denoted T*ML_{dpx}*F), including replicate curves at each of the 11 conditions of temperature (15–30°C) and [TBP] (200–800 nM). Each of the resulting 11 averaged curves was best described by triexponential decay of the form: $Y = \alpha_1 \exp(-\lambda_1 t) + \alpha_2 \exp(-\lambda_2 t) + \alpha_3 \exp(-\lambda_3 t)$. It was clear from both these and previous measurements and analyses [Parkhurst et al 1996]

that the slowest decay process was attributable to fluorescein photobleaching and does not derive from the DNA-TBP interaction; this phase of the curve was therefore excluded from all subsequent analyses. The resulting set of biexponential decay curves showed an amplitude change of 44(±1.4)% in fluorescein emission upon binding of TBP to T*ML_{dpx}*F with no apparent pattern based on temperature or [TBP]. These results are identical within experimental error to the steady-state change in fluorescence intensity independently measured at each temperature (Figure 1). These observations confirm that ascribing the third kinetic phase to photobleaching does not eliminate a TBP-DNA binding phase.

The biphasic character of the binding kinetics decreases with increasing temperature and protein concentration to nearly monophasic at 30°C and high [TBP] (Figure 2; [Parkhurst et al 1996]). The eigenvalues associated with the observed biexponential decays were non-linear with [TBP]. The amplitudes and eigenvalues were used only to construct smoothed representations of the observed decay curves, in the form of 41 points/curve, for global analysis, but were themselves not explicitly fit (see Experimental Procedures).

Equilibrium and relaxation constants obtained by steady-state FRET

Figure 3 shows the equilibrium titration data obtained using FRET as well as the fitted curve for the binding of the T*ML_{dpx}*F to TBP at 20°C. The K_a values of 29.2(±5.3) (15°C), 48.2(±5.6) (20°C), 61.7(±6.6) (22.5°C), 78.5(±9.4) (25°C), and 125.0(±22) (30°C) μM^{-1} were determined by van't Hoff analysis (Figure 3, inset).*

A relaxation constant, R, is a measure of the rate of DNA replacement from DNA-TBP complexes and reflects dissociation of DNA from all bound species. The dominant slow phase of the TBP-AdMLP relaxation is temperature independent. The relaxation data at all five temperatures were well described by a biexponential decay with $R = 1.7(\pm 0.2)\times 10^{-3} s^{-1}$(Figure 4). This slow phase comprised ~90% of the total change in amplitude, depending on the temperature, and only this well-determined relaxation constant for the slow phase was used in the analysis. The small initial transient phase (~0.1 s^{-1}) was too rapid for precise determination in the hand-mixing experiments. This transient rapid phase would have been used in the global analysis only to reject any fits that failed to find that phase; due to the quality and quantity of the data included in the analysis, that additional small constraint was found to be unnecessary.

* Analysis of these data shows that binding over the temperature range of these FRET measurements (15–25°C) is described by a linear regression within the error of the measurements. Equilibrium binding studies conducted over a wider temperature range are consistent with a negative ΔC°_p for the binding of TBP to the major late promoter [Petri et al 1998].

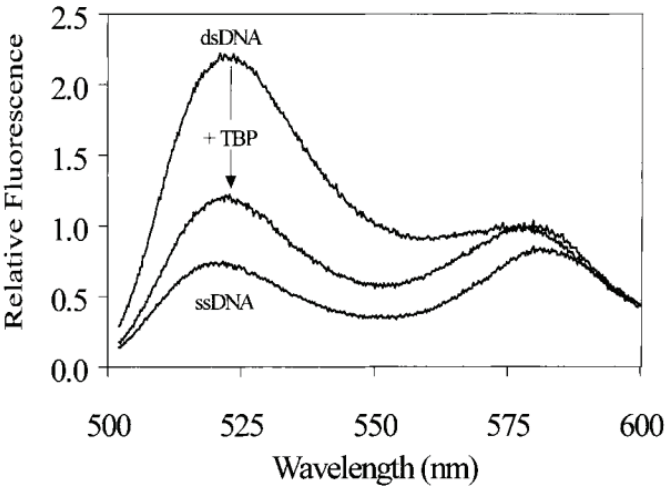

Figure 1. Steady-state fluorescence emission spectra of 4 nM T*ML_{dpx}*F as a single strand (ssDNA) and in a duplex both free in solution (dsDNA) and bound to TBP at 20°C. The 520 nm and 580 nm fluorescence emission peaks correspond to fluorescein and TAMRA emission, respectively. Upon addition of complement, duplex formation increases the distance between the two fluorophores, decreasing the resonant transfer of energy. (Changes in the 580 nm region reflect changes both in FRET and in static quenching of TAMRA.) The addition of TBP results in DNA bending, a decrease in the distance between the fluorophores and an increase in FRET. Neither the spectrum of free nor TBP-bound T*ML_{dpx}*F changes significantly with temperature from 15–30°C. (The donor-only duplex spectrum changes 1.9(±1.6)% upon addition of TBP, independent of temperature.)

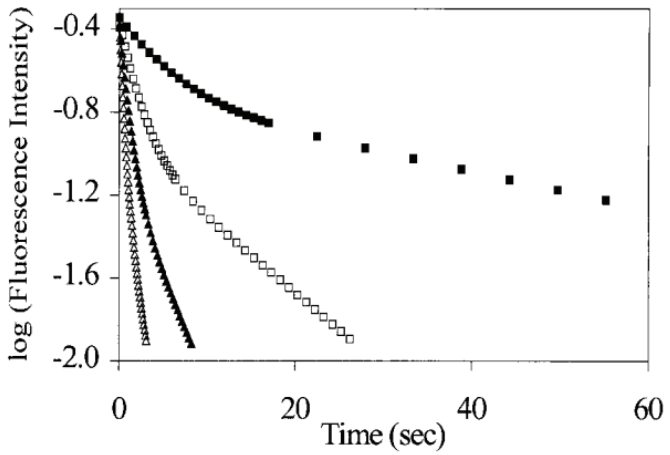

Figure 2. Stopped-flow association kinetic curves for 20 nM T*ML_{dpx}*F reacting with 400 nM TBP at 15°C (■), 20°C (□), 25°C (▲) and 30°C (Δ). The biexponential decays corresponding to the observed progress curves are shown. The values observed for λ_1 are 0.20, 0.54, 1.1 and 2.5 s^{-1}, respectively.

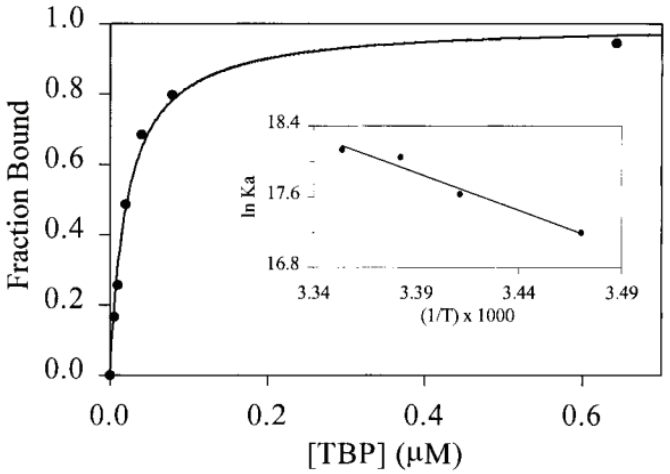

Figure 3. Equilibrium binding isotherm for AdMLP with TBP at 20°C. The fraction of bound TBP at each titration point is calculated directly from the change in the steady-state emission spectrum. The van't Hoff plot shown in the inset is constructed from the experimentally derived K_a values and yields a value for $\Delta H°$ of 16.9(±2.2) kcal/mol.

Global analysis

Global analysis was conducted to determine the correspondence of the complete set of kinetic and thermodynamic data to various kinetic models. The identification and quantification of intermediates in the TBP-DNA reaction pathway is critically dependent upon the use of global analysis procedures in order to reach a unique solution of the microscopic rate equations that characterize a given mechanism. The agreement between experiment and theory for each model was evaluated by

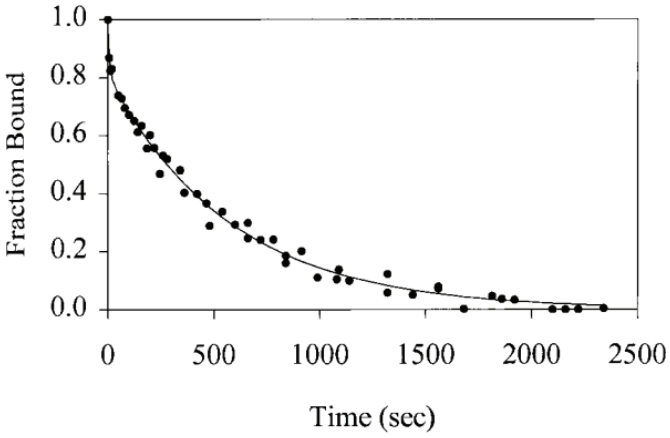

Figure 4. Relaxation kinetics measured at 30°C determined from changes in steady-state fluorescence emission. The time-course of the release of T*ML_{dpx}*F from the TBP-bound complex was monitored following addition of a large excess of unlabeled DNA. The initial fast phase of the relaxation reflects the direct dissociation of $I_1 \rightarrow$ TBP+DNA from the equilibrium population of TBP-DNA. The amplitude and rate constant for this phase correspond within error to those obtained from simulations using values obtained from global analysis.

the consistency of the kinetic and thermodynamic parameters determined from the analysis with the combined experimental data. The overall sum of squared residuals, σ_{global}^2, was determined for all models according to the expression:

$$\sigma_{global}^2 = \frac{1}{5}\left[\frac{3}{2}\left(\sigma_{curve}^2 + \left(\frac{\sigma_\lambda^2 + \sigma_{amp}^2}{2}\right)\right) + \sigma_K^2 + \sigma_R^2\right] \quad (1)$$

where σ_{curve}^2, σ_λ^2 and σ_{amp}^2 reflect the quality of the fit to the stopped-flow kinetic curves, and σ_K^2 and σ_R^2 reflect the quality of the fit to the equilibrium and relaxation data, respectively. An ideal global fit to a model would yield values for all terms that contribute to the sum of squared residuals that are less than or equal to their individual respective errors, resulting in $\sigma_{global} \leqslant 1.0$.

In these analyses, the rate constants were assumed to be described by the usual transition-state model. All of the fitting parameters, including the three relative quantum yields, were globally linked in the terminology used by [Knutson et al 1983]. The assumed temperature independence of the relative quantum yields for the intermediate species followed directly from the observation that quantum yields for both the initial state (free T*ML_{dpx}*F) and the final state (TBP-bound T*ML_{dpx}*F) are independent of temperature over the range of these measurements. It was deemed unlikely that the rate of internal conversion or intersystem crossing would vary over this range for a given intermediate species. Variation of the quantum yields of intermediates with temperature could conceivably derive from equilibria with a non-zero $\Delta H°$ between two forms of a given species of labeled DNA, with the two forms having different quantum yields. This added level of complexity, which would result in six additional parameters for the two-intermediate model and an indefensible analysis, was found to be unnecessary.

The correspondence of the stopped-flow kinetic, equilibrium and relaxation data was first tested against a one-intermediate model:

$$\text{TBP} + \text{DNA} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \underset{\text{intermediate}}{\text{TBP}\cdots\text{DNA}} \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \text{TBP-DNA}_{\text{FINAL}} \quad (2)$$

By themselves, each of the 11 stopped-flow progress curves are well described by this model. In this case, only σ_{curve}^2 contributes to the goodness of the fit and $\sigma_{global} = 0.947$. The average residual is 0.018, within the average error of the individual data points (see Experimental Procedures). Inclusion in the global fit of the eigenvalues and amplitude changes increases σ_{global} to 1.08 (which derives from σ_{curve}^2, σ_λ^2 and σ_{amp}^2) with the average residual still within error. The calculated values of $(\alpha_1+\alpha_2)$ were within the 95% confidence region, as were all but one of the calculated values for λ_1. Only the first eigenvalue for each curve was used in assessing the overall qual-

ity of the fit in the global analysis, since its value was significantly better determined experimentally than was the value of λ_2.

Inclusion of the relaxation and equilibrium binding data* in the global analysis resulted in a severe lack of fit to the single-intermediate model. The best-fit value of σ_{global} was 6.04 and the average residual for the stopped-flow curves was 0.089. All fitted values of λ_1 and $(\alpha_1+\alpha_2)$ were outside the 95% confidence region. The calculated values of K_a ranged from 76.6–80.4 μM^{-1}, compared to the observed values range of 29.2–125 μM^{-1}, with four of the calculated K_a values outside the 95% confidence limit. The best-fit values of the relaxation constant, R, ranged from $0.33 \times 10^{-3} s^{-1}$ to 4.10×10^{-3} s^{-1}(compared to the measured value of $1.7(\pm 0.2) \times 10^{-3}$ s^{-1}) and three of the fitted values were outside the 95% confidence region. The best fit to the stopped-flow curves obtained fitting all data to the one-intermediate model is shown in Figure 5(a). Clearly the single-intermediate model (equation (2) is insufficient to describe the complete ensemble of equilibrium and kinetic data.

The data were next fit to a two-intermediate model:

$$\begin{aligned} \mathrm{TBP} + \mathrm{DNA} &\underset{k_2}{\overset{k_1}{\rightleftharpoons}} (\mathrm{TBP}\cdots\mathrm{DNA})_1 \\ &\underset{k_4}{\overset{k_3}{\rightleftharpoons}} (\mathrm{TBP}\cdots\mathrm{DNA})_2 \\ &\underset{k_6}{\overset{k_5}{\rightleftharpoons}} \mathrm{TBP\text{-}DNA_{FINAL}} \end{aligned} \qquad (3)$$

where I_1 and I_2 are the first and second intermediates, respectively. K_a for this model is defined as $(I_1+I_2+\text{final complex})/[(\text{DNA})(\text{TBP})]$ or, in terms of the microscopic rate constants:

$$K_a = \frac{k_1}{k_2}\left(1+\frac{k_3}{k_4}\left(1+\frac{k_5}{k_6}\right)\right) \qquad (4)$$

and $R = k_2k_4k_6/(k_3k_5 + k_2k_5 + k_2k_4)$. Fitting the stopped-flow, equilibrium and relaxation data to the two-intermediate model (equation (3) yields $\sigma_{global} = 0.995$ and an average residual to the points describing the stopped-flow curves of 0.023, the latter value being identical with the average measured residual. All calculated eigenvalues and amplitude changes for the stopped-flow curves were within the 95% confidence region, as were all calculated values for the equilibrium and relaxation constants. Table 1 shows the optimal fitted values obtained for each of the six rate constants at 25°C and their associated enthalpies of activation. The errors associated with each of these values were determined as described in Experimental Procedures. A best-fit curve calculated for the two-intermediate model and the stopped-flow data is shown in Figure 5(b).

The calculated mole fraction of each species as the reaction progresses is shown in Figure 6 for (a) 15°C and (b) 30°C. The relative proportions of these species vary with TBP concentration as well as with temperature. The first intermediate, I_1, is present at significant mole fractions at all temperatures and [TBP], with the maximal mole fractions for I_1 ranging from 0.66 at 15°C to 0.92 at 30°C. In contrast, the second intermediate, I_2, is present at relatively low concentrations with maximal mole fractions ranging from only 0.097 at 15°C to 0.014 at 30°C. It should be noted that at 30°C, I_1 is present at high concentration for a significant proportion of the reaction time-scale.

Analyses consistently found multiple bound forms of the TBP-DNA complex to be present at equilibrium even with saturating [TBP]. At equilibrium at 30°C and 800 nM TBP, I_1 comprises 16.5% of the species present in solution, with 82.1% as final complex and the remaining 1.4% as I_2 and free $T^*ML_{dpx}^*F$. The fast phase of the relaxation (Figure 4) derives essentially from $I_1 \rightarrow$TBP+DNA, with the amplitude corresponding to the mole fraction of I_1 at equilibrium and the rate $\approx k_2$. The quantum yield ratios for I_1, I_2 and the final complex, relative to a defined value of 1 for the free $T^*ML_{dpx}^*F$,

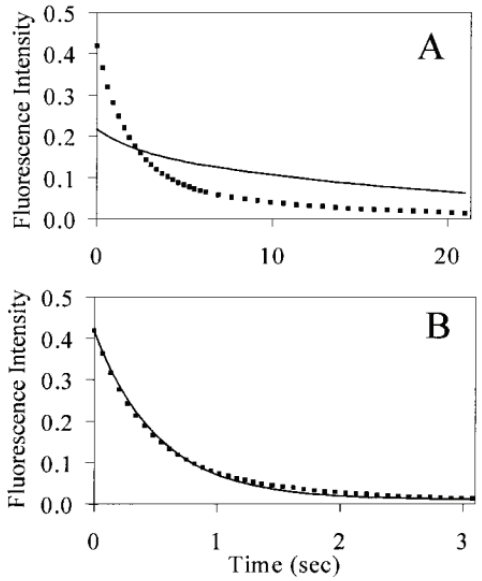

Figure 5. (a) Best global fit to the stopped-flow curves obtained using the one-intermediate model, with [TBP] = 800 nM and 20°C. (b) Best global fit to the stopped-flow curves obtained using the two-intermediate model, with [TBP] = 400 nM and 30°C. The sum of squared residuals for this curve is <0.1 that of the curve shown in (a). The associated value of λ_1 is 2.5 s^{-1}. For [TBP] = 200 and 800 nM, λ_1 = 2.1 and 3.6 s^{-1}, respectively.

* The expressions for K_a and R, in terms of the microscopic rate constants associated with the one-intermediate model are $(k_1/k_2)(1+k_3/k_4)$ and $k_2k_4/(k_2+k_3)$, respectively.

Table 1. Globally derived kinetic and thermodynamic parameter values corresponding to the two-intermediate model (equation (3)

i	k_i (s^{-1})	$\Delta H_i^{\circ\neq}$ (kcal/mol)	$\Delta S_i^{\circ\neq}$ (cal deg^{-1} mol^{-1})	$\Delta G_i^{\circ\neq}$ (kcal/mol)
1	1.59 (1.56,1.66)	35.1 (34.7,35.5)	87.8 (86.4,89.1)	8.990 (8.966,9.002)
2	0.118 (0.114,0.125)	18.1 (16.9,18.7)	−2.19 (−5.96,0.0693)	18.71 (18.68,18.74)
3	0.0288 (0.0283,0.0339)	3.02 (3.02,4.35)	−55.4 (−55.4, − 50.8)	19.55 (19.46,19.56)
4	0.541 (0.534,1.01)	29.3 (28.5,30.1)	38.6 (36.5,41.9)	17.82 (17.45,17.82)
5	0.365 (0.365,0.730)	26.7 (25.1,27.1)	29.2 (24.5,30.9)	18.04 (17.65,18.05)
6	0.00354 (0.00349,0.00395)	4.02 (3.98,4.75)	−56.3 (−56.4, − 53.7)	20.79 (20.73,20.80)

Values shown are for 25°C and 1 M standard state for each "step", i, along the reaction pathway (see Figure 7). Numbers in parentheses represent the upper and lower error bounds corresponding to the 68% confidence region, determined as described in Experimental Procedures. k_1 is a second-order rate constant with units of $\mu M^{-1}s^{-1}$;k_2 through k_6 are first-order rate constants. Because I_1 and I_2 as well as the final complex are present at equilibrium for infinite [TBP], the measured $\Delta H^\circ = \Sigma\, \varphi_i \bar{H}^\circ_i$, where i indexes the three bound species, φ is mole fraction and $\bar{H}$ is the mole enthalpy referenced to $\bar{H}^\circ = 0$ for free DNA. For 25°C, this apparent ΔH° = 13.4 (11.6,15.2) kcal/mol.

were 0.547 (0.539, 0.549)*, 0.574 (0.562, 0.605), and 0.581 (0.575, 0.582), respectively. These values indicate that the distance between the donor and acceptor molecules and thus the inferred degree of DNA bending is essentially the same in the intermediate and final complexes.

The thermodynamic profile of the reaction pathway

The thermodynamic profile for this TBP-DNA association, including ΔH°, ΔS° and ΔG° for each step along the reaction pathway, is shown for a 1 M standard state in Figure 7. The overall free energy change is −10.7 (−10.8, −10.7) kcal/mol. The reaction is entropy-driven under these experimental conditions with ΔS° = 81.3 (76.6, 86.6) cal/K mol. The largest energetic barrier in the reaction is the initial binding step which has an activation enthalpy of 35.1 (34.7, 35.5) kcal/mol. The largest increase in entropy also occurs at this step. The largest decrease in ΔG° occurs upon the initial binding/bending event with the formation of I_1. The second reaction step is enthalpy-driven and is accompanied by a large decrease in entropy. In overall free energy, I_2 differs from I_1 by +2.0 kcal/mol. The third step is again entropically driven, with a significant energy barrier overcome in the formation of the final complex. The final complex is 0.9 kcal/mol lower in ΔG° than I_1.

Stoichiometric binding studies

Analytical ultracentrifugation studies have shown that full-length *S. cerevisiae* TBP is monomeric under the conditions and protein concentrations used in these studies [Daugherty et al 1999]. As independent confirmation that TBP oligomerization does not impact upon the studies described here, kinetics experiments were conducted with a 1:1 ratio of DNA and TBP. When TBP was diluted directly from its cold 43.9 μM stock solution into five DNA solutions at equimolar concentrations of 0.04, 0.5, 1.5, 4 and 8 μM, identical 42.4(±2.8)% changes in fluorescein emission were observed within 20 seconds of TBP addition. The emission spectra were constant for the additional 30 minutes that the reactions was monitored. Identical results were obtained when the labeled $T^*ML_{dpx}{}^*F$ probe comprised 5, 10 or 100% of the total DNA (with the balance of the DNA being unlabeled ML_{dpx}), showing that $T^*ML_{dpx}{}^*F$ is a reliable trace probe for binding studies conducted at high concentrations of DNA.

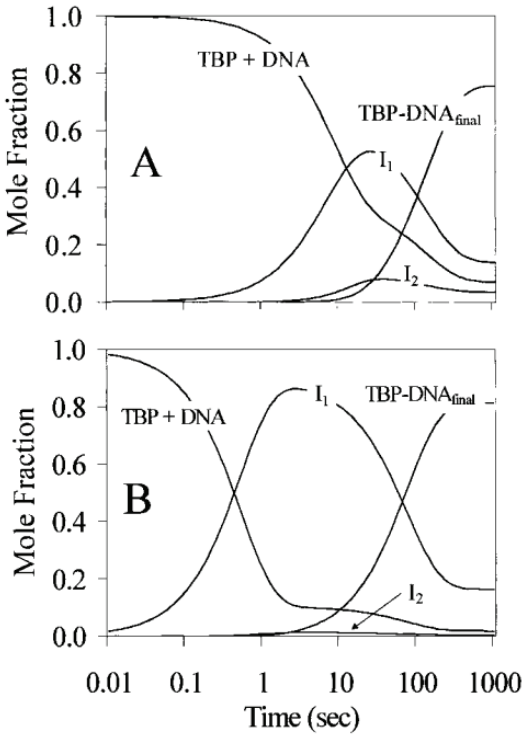

Figure 6. The mole fraction of each species along the reaction pathway at (a) 15°C and (b) 30°C, generated from the globally derived parameters with [TBP] = 400 nM. The dominance of I_1 at 30°C over the relevant time interval is apparent. The mole fraction of each species varies with both temperature, as shown here, and with [TBP].

* Values in parentheses indicate the upper and lower error estimates that were determined as described in Experimental Procedures.

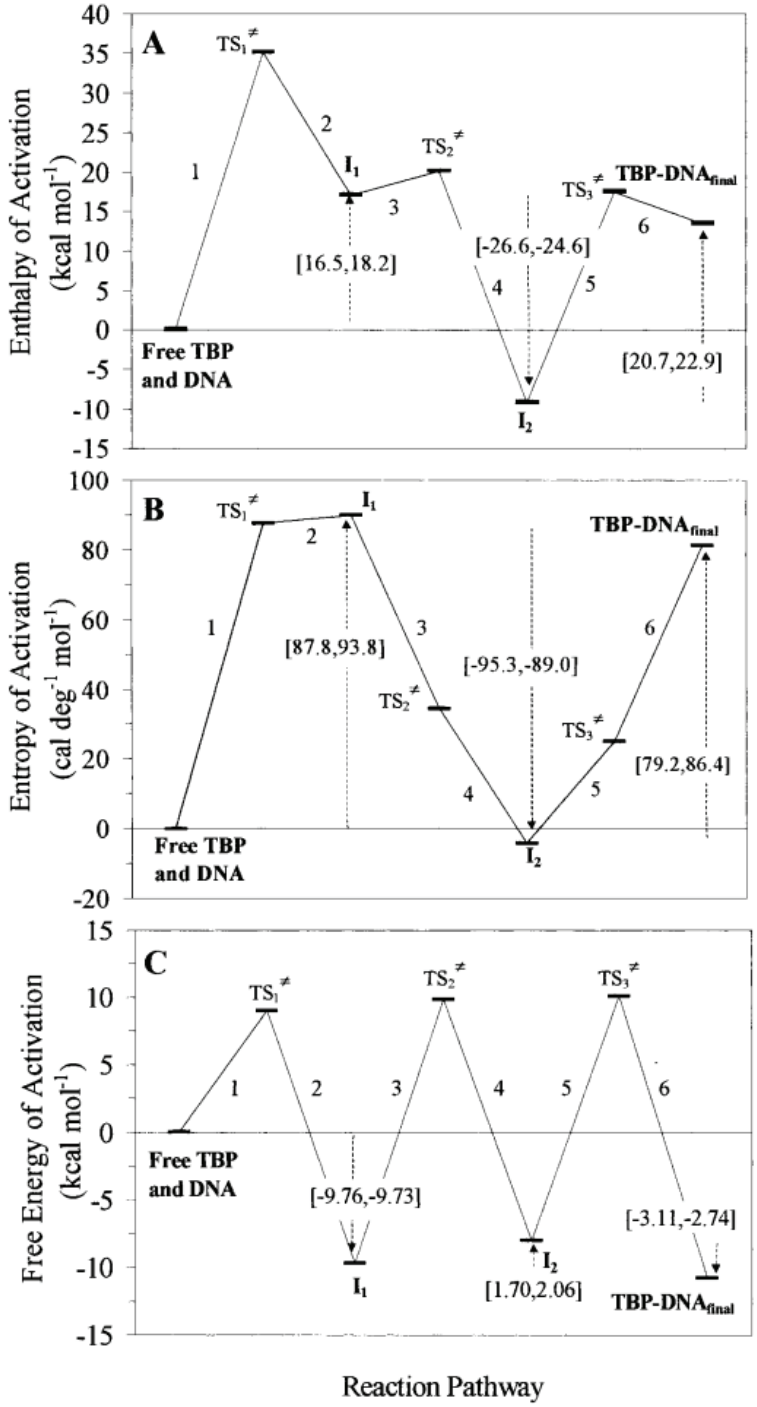

Figure 7. A graphical representation of the thermodynamic changes associated with each step along the two-intermediate reaction pathway at 25°C and 1 M standard state for (a) enthalpy of activation, (b) entropy of activation and (c) free energy of activation. Numbers 1 through 6 correspond to the respective steps, *i*, in Table 1. The transition states are denoted by $TS_i^{\neq}$. The changes in thermodynamic values between successive species are indicated by the broken arrows, with the upper and lower error limits for each change bracketed.

Discussion

The complex kinetics observed by FRET reflect the interaction of monomeric TBP with the major late promoter

Analytical ultracentrifugation studies have demonstrated that, for [TBP] $\leqslant$1.5 μM, the *S. cerevisiae* protein is $\geqslant$90% monomer under the experimental conditions of the studies presented here [Daugherty et al 1999]. Thus, TBP is almost exclusively monomeric in these stopped-flow studies, since its maximum concentration was 1.6 μM prior to mixing. Nevertheless, we directly verified that TBP oligomerization has not influenced the results and conclusions drawn from these FRET studies. If a slow self-association equilibrium were present that was competitive with DNA binding, and DNA bound only to monomeric TBP (as has been reported for human TBP; [Coleman et al 1995 and Coleman and Pugh 1997]), the addition of TBP to $T^*ML_{dpx}^*F$ at a 1:1 concentration ratio would yield an initial fast binding event with a small amplitude reflecting the low concen-

tration of monomer. The remainder of the total amplitude change would then appear slowly as oligomeric TBP dissociated into monomer and bound DNA. In contrast, we observed that when equimolar concentrations of DNA and TBP were mixed at concentrations up to 8 μM, 100% of the total expected amplitude change occurred within 20 seconds, directly demonstrating that such a protein oligomerization equilibrium does not compete with DNA binding under our solution conditions. This result directly contradicts the incorrect assertion by Pugh and co-workers [Coleman and Pugh 1997], that such a linked reaction, with very slow dissociation of dimers, explains our previously published studies [Petri et al 1995 and Parkhurst et al 1996]. Thus, independent lines of experimental evidence demonstrate that the observed kinetic complexity of the DNA binding reaction does not derive from TBP oligomerization. Conversely, the FRET experiments conducted with equimolar 8 μM DNA and *S. cerevisiae* TBP (where the protein is 95% octameric; [Daugherty et al 1999]) are consistent with either binding of DNA to oligomeric TBP or equilibration of TBP oligomers that is very fast relative to DNA binding.

Kinetic mechanisms of TBP-DNA interactions

Determination of the kinetic mechanism(s) of sequence-specific TBP-DNA association is a key step towards understanding the initiation of gene transcription and its regulation. Hawley and co-workers established the second-order nature of TBP-DNA complex formation [Hoopes et al 1992], and also postulated a "two-step" mechanism. Their proposal that an intermediate was present was based on the observation that the rate constant determined for the association process was significantly slower than diffusion limited ($2.6 \times 10^5 M^{-1}s^{-1}$) and on the limiting values inferred from a double-reciprocal plot of the rate data. Formation of this intermediate was understood to precede the association process that was being measured, with the postulated intermediate complex loosely bound. Stopped-flow DNase I footprinting kinetics studies conducted at higher concentrations of TBP were similarly analyzed and interpreted to suggest an even more unfavorable initial equilibrium and a faster second step in this mechanism [Petri et al 1995]. However, the fact that the intercepts of the aforementioned double-reciprocal plots are zero within experimental error is appropriately interpreted as the postulated intermediate being simply a transient "encounter complex". This precludes subsequent assignment of values to various constants in a multiple-step model.

Inclusion of a transient "encounter complex" in the reaction mechanism

A transient encounter complex is tacitly understood to exist in all solution reactions and is not ordinarily shown in a reaction mechanism. The encounter complex was explicitly included in the discussion and interpretation of our original FRET studies [Parkhurst et al 1996]. For consistency, the one and two-intermediate models analyzed here are shown below to include the encounter complex:

$$\mathrm{TBP} + \mathrm{DNA} \overset{K_{eq}}{\rightleftharpoons} \underset{\text{encounter complex}}{(\mathrm{TBP}\cdots\mathrm{DNA})} \overset{k_i}{\underset{k_2}{\rightleftharpoons}} \mathrm{I}_1 \overset{k_3}{\underset{k_4}{\rightleftharpoons}} \mathrm{TBP\text{-}DNA_{FINAL}} \quad (5)$$

$$\mathrm{TBP} + \mathrm{DNA} \overset{K_{eq}}{\rightleftharpoons} \underset{\text{encounter complex}}{(\mathrm{TBP}\cdots\mathrm{DNA})} \overset{k_i}{\underset{k_2}{\rightleftharpoons}} \mathrm{I}_1 \overset{k_3}{\underset{k_4}{\rightleftharpoons}} \mathrm{I}_2 \overset{k_5}{\underset{k_6}{\rightleftharpoons}} \mathrm{TBP\text{-}DNA_{FINAL}} \quad (6)$$

The model represented in equation (5) cannot accommodate the kinetic and thermodynamic data. Such a single-intermediate model has been proposed [Hoopes et al 1998]. The rate constants associated with formation and dissociation of the encounter complex are not known. However, if both steps are rapid, formation of the encounter complex can be represented by an equilibrium constant. Equations (2) and (5) for the one-intermediate model and (3) and (6) for the two-intermediate model are then related by $k_1 = K_{eq}k_i$ with the overall equilibrium constants, K_a, equal to k_1k_3/k_2k_4 and $k_1k_3k_5/k_2k_4k_6$, respectively.

TBP binding and DNA bending are not separable

In our initial FRET studies of promoter-TBP binding conducted at 30°C, simple second-order kinetics were observed, although a subtle lack of fit in the data analyses was noted [Parkhurst et al 1996]. From the work reported here, it is clear that the multiphasic nature of the kinetics of this interaction is highly temperature dependent. It was by extending the kinetic studies to a range of temperatures and [TBP] and including both equilibrium and relaxation data that the sequential two-step mechanism for TBP binding to the major late promoter was revealed.

Contrary to many investigators' preconception, TBP binding and DNA bending are not separable processes. Even in the complex reaction pathway that has been elucidated, TBP binding and DNA bending occur in concert. This conclusion is based on the observation that the relative quantum yields of the intermediate species and the final complex are nearly equivalent, reflecting similar DNA bending among the three TBP-DNA complexes.

A two-intermediate model describes this TBP-DNA interaction

The multiphasic character of the kinetic curves, particularly at lower temperatures and [TBP], show that the TBP-DNA reaction mechanism must have more than a single step. (For any sequential reaction with a reversible first step, the observed decay would be multiphasic even in the case where all bound forms of DNA had identical quantum yields.) Four additional observations are consistent with the presence of at least one intermediate. First, R is so small that it cannot influence the forward rate measured by λ_1. Consequently, were there no intermediates, λ_1 would be a simple pseudo-first-order forward rate constant linear in [TBP] and would approach zero as [TBP] approaches zero, in contrast to observation. Second, the ratios of the kinetic constants for the dominant phases of the forward and reverse processes (λ_1 and R, respectively) differ from the measured equilibrium constants by factors of 10–30. Third, the difference in the activation energies for these processes (ca 30 kcal/mol) is in marked disagreement with the measured $\Delta H°$ of 16.9 kcal/mol. Fourth, the small rapid component observed in the reverse reaction (Figure 4) can derive only from two or more bound forms of TBP-DNA at equilibrium.

The global analysis of the association and relaxation kinetics and equilibrium binding data unequivocally disproved the simplest one-intermediate model. We further considered whether bi-directional TBP binding was the source of the two kinetic phases, since *S. cerevisiae* TBP can bind TATA box sequences in both of the two possible orientations with ~3:2 distribution ratio at equilibrium [Cox et al 1997]. If this bi-directional binding were responsible for the two observed kinetic phases, the simplest associated kinetic model would be a two-step branching mechanism with a single step leading from the reactants to each final complex. (This model would also describe heterogeneous TBP reacting in a single step.) Given the tight binding and slow dissociation of the TBP-TATA complex, the ratio of the two amplitudes would be invariant at the 3:2 ratio and the values of λ_1 and λ_2 would be linear with [TBP] if this model were applicable. That this prediction is not consistent with the experimental results shows that bi-directional binding of TBP is not the source of the observed multiple kinetic phases.

Failure of the one-intermediate and two-step branching models to accommodate all of the experimental data necessitated consideration of the next, more complex model, a sequential, two-intermediate process. This model successfully describes all of the experimental data. At any single temperature, the amplitudes and eigenvalues of the two phases in the stopped-flow kinetic curves contain sufficient information as a function of [TBP] to determine three of the rate constants and two of the quantum yield ratios. (Although three non-zero eigenvalues may be derived from this model, λ_2 and λ_3 could not be separately extracted from the data.) A fourth rate constant is determined from K_{eq}. Including the relaxation constant, R, allows determination of a fifth rate constant.

Extension of the data analysis to two temperatures provides five activation enthalpies in addition to the five rate constants and two relative quantum yields noted above. Since each rate constant was constrained to a transition-state two-parameter model, data at only three temperatures is sufficient to constrain the kinetic parameters and provide confidence intervals for the parameters. Further extension to a global analysis that encompasses the temperature and [TBP] dependence of the kinetic course of the reaction, allows extraction from the data of the sixth rate constant, its activation enthalpy and the third relative quantum yield. The ability of global analysis procedures to extract precise model-dependent parameter values is well documented [Knutson et al 1983]. Thus, the combined association kinetic, equilibrium and relaxation data contain sufficient information to allow extraction of the 15 parameters completely characterizing this model. These parameters include the six rate constants at 30°C, six activation enthalpies and three relative quantum yields. That this kinetic model is well determined by the data is indicated by the fact that the global analysis converged to essentially the same parameter values when starting from many very different initial values and the relatively small errors obtained for each parameter. Under the conditions of the measurements, the concentration of the second intermediate species is always relatively very low and the data contain the least information about this species. As would be expected, the errors in the rate constants leading away from this species and the associated activation enthalpies are the least well determined. On the other hand, inclusion of this second intermediate is essential for accommodating the complete ensemble of data. Thus, the sequential two-intermediate process is the simplest model consistent with the data.

Relating the microscopic rate constants determined using FRET to the steady-state processes measured by mobility-shift and footprinting

For a complex linear first-order mechanism, it is possible that only a single eigenvalue may be observed for either association or dissociation. These eigenvalues correspond to a quasi steady-state approach to equilibrium in the forward and reverse directions, respectively. Such values would be expected from measurements of the time-course of a reaction that encompass primarily the steady-state (slow) phase and not the initial (rapid) phase of the reaction. These forward and reverse steady-state decay constants, R_f and R_r, respectively, are expressed as combinations of the microscopic rate constants. To the extent that the time derivatives of the intermediate species

are equal to zero (the steady-state hypothesis), R_f+R_r will be the first-order steady-state decay constant for the approach to equilibrium from either direction.

The time-course of the association process for DNA-TBP complex formation is composed of a rapid transient phase during which I_1 is built up followed by a slow, quasi steady-state process during which I_1 remains nearly constant and at high mole fraction. For the two-intermediate mechanism, to the extent that the slowest process approximates a steady-state process, $R_f = k_1$ (TBP) $k_3k_5/(k_3k_5 + k_2k_5 + k_2k_4)$*. At 30°C, R_f/[TBP] = 3×10^5 $M^{-1}s^{-1}$, in very good agreement with the association rate constants determined from studies conducted by gel mobility-shift [Hoopes et al 1992] and DNase I footprinting [Petri et al 1995 and Petri et al 1998]. A value in this range has also been reported from a fluorescence anisotropy study [Perez-Howard et al 1995]. In addition, the activation energy calculated for R_f agrees very well with that determined from footprinting over this temperature range [Petri et al 1998].

For the relaxation experiments, $k_1 \approx 0$ and the relaxation constant (R) $\approx R_r$. Simulations of TBP-AdMLP complex relaxation performed using the two-intermediate model and the globally determined rate constants show that a small and fast transient phase is followed by a slow steady-state phase with negligible intermediate populations, consistent with the steady-state hypothesis.

Finally, to the extent that the slower phase of association approximates a process described by R_f, the value of K_{eq} can be determined from R_f/R_r, and the values of K_{eq} determined in the above-cited studies agree very well with the values determined here. Taken together, these observations lend support to the conclusion that these other experimental techniques are primarily measuring steady-state processes, and highlights their general consistency with the FRET studies presented here. It should be noted that the extent to which any system can be accurately described by R_f and R_r is apparent only in retrospect, after the microscopic k values have been determined.

*The time interval over which I_1 is within 25% of its peak value corresponds to the interval encompassing 10–90% of a steady-state process calculated from R_f for [TBP] = 200–800 nM at 30°C. R_r is negligible for the association reactions in this study.

† Both entropy and free energy changes for bi-molecular processes, such as DNA+TBP→I_1, are dependent on the choice of standard state [Gurney 1953 and Cantor and Schimmel 1980]. A standard state of one mole fraction converts the entropy (and free energy) changes to unitary changes, thereby removing the cratic contribution, R (ln 55) and allowing $\Delta S°_u$ and $\Delta G°_u$ for bi-molecular and uni-molecular changes to be compared. Conversion results in a decrease of eight entropy units for the initial step, NA+TBP→I_1; all values of $\Delta S°$ for uni-molecular steps are unchanged. Correspondingly, $\Delta G°_u$ is 2.4 kcal/mol more positive than $\Delta G°$.

What are the structures of I_1 and I_2?

While the structures of I_1 and I_2 are clearly of great interest, the only direct insight into them at present derives from the fitted values of their relative quantum yields. We have previously shown the consistency of the distance distribution for TBP-bound T*ML_{dpx}*F, obtained from FRET lifetime measurements, with the bent DNA in the co-crystal structure [Parkhurst et al 1996]. The relative quantum yield obtained in this study for the final complex must necessarily be consistent with DNA bent to that extent. The equivalence of the relative quantum yields for the bound forms suggests that I_1, I_2 and the final complex have similarly bent DNA. A caveat to this proposal is that FRET measurements are sensitive only to the end-to-end distance between the donor and acceptor probes, rather than to the specific conformations of the molecules. It is possible, although we deem it unlikely, that non-dipolar effects unique to the intermediate species are responsible for the I_1 and I_2 quantum yields, or that different DNA-bending (relative to the final complex) in I_1 and I_2 is exactly compensated by some non-dipolar change in the intermediates' quantum yields. The simplest interpretation of the nearly equivalent quantum yields is that the DNA in the intermediate and final complexes is nearly equivalently bent, in which case these species most likely differ in the details of the protein-DNA interface.

Endothermic changes of equally large magnitude are associated with both the initial, (DNA+TBP→I_1) and final (I_2→final complex) steps in the reaction (Figure 7(a) with the commensurate entropic changes, after conversion to a unitary standard state,† identical within error. Linking these two thermodynamically similar processes is the formation of I_2 from I_1, which is highly exothermic and accompanied by a large decrease in entropy. We have suggested [Parkhurst et al 1996] that the highly endothermic nature of the initial reaction step reflects the enthalpy change associated with an equilibrium between the free DNA and DNA "pre-bent" to a conformation compatible with the concave surface of TBP. Thus, for instance, I_1 could represent a TBP-DNA complex in which neither pair of phenylalanine residues has intercalated into the DNA "kinks", I_2 a complex with one intercalation event with the final complex possessing the structure observed by crystallography. An important goal of future experiments is to determine the structural correlates associated with each of these thermodynamic changes.

Do the TBP-TATA intermediate species mediate assembly of the transcription pre-initiation complex?

The identification of an intermediate that is energetically stable and has bent DNA raises the possibility that pre-initiation complexes (PICs) could form by reaction of subsequent tran-

scription proteins with I_1, the species that dominates much of the TBP association reaction but is not the thermodynamically most favored form. With assembly of PIC from I_1 rather than from the final complex, the reaction would be under kinetic rather than thermodynamic control. The simplest scenario of the TBP-TATA complex binding to one RNA polymerase-specific factor, TFIIB, has been explored as a first-approximation of transcription initiation under kinetic rather than thermodynamic control.

The possible biological consequences of kinetic control of PIC formation are clearest when the two-intermediate model is considered together with bi-directional binding of TBP, where TBP binds with a very slight preference for one orientation, both TBP-DNA orientations bind TFIIB [Cox et al 1997] and transcribing complexes require the C-terminal end of TBP to be oriented toward TFIIB and toward the 3′ end of the coding strand. The "orientational commitment" of the TBP is assumed to occur after ternary complex formation so that the subsequent events resulting in PIC formation are irreversible. Simulations for this model were conducted, with the TBP-DNA association steps of both orientational pathways assigned the rate constants obtained from the global analysis, since the free energy for the final complex in either orientation is nearly identical [Cox et al 1997]. An association rate constant of 3×10^6 M^{-1} s^{-1} was assumed for the binding of TFIIB to the productive TBP-DNA complex, and the concentrations of TBP, DNA and TFIIB were each assumed to be 10 μM [Lee and Young 1998 and Sethy-Coraci et al 1998]. (Varying TBP, TFIIB and DNA from 0.5–10 μM and the rate constant for TFIIB binding from 0.1×10^6 to $300\times 10^6 M^{-1}$ s^{-1} yielded comparable simulated results.) The ratio of the rate constants for formation of productive: non-productive ternary complexes was assumed to be 4:1, reflecting the observed orientational distribution [Cox et al 1997].

With thermodynamic control of PIC formation, TFIIB binds only to the final TBP-DNA complex, whereas with kinetic control of this process TFIIB binds predominantly to I_1. In both cases, only one binding orientation can lead to PIC formation. With thermodynamic control of this process, six minutes are required after the start of the reaction for ~50% of the TBP to be directly incorporated into correctly oriented ternary complexes that proceed along the initiation pathway. The remaining ~50% of the TBP is incorporated by 0.25 second into transcriptionally non-productive, slowly dissociating complexes ($t_{1/2}$ = 44 minutes), with > 90% transcription occurring only 1.5 hours after the initiation of protein binding. The approach to full transcription induction is very slow due to the kinetic trapping of the proteins in the incorrect orientation. In distinct contrast, with kinetic control only 0.5 second are required for incorporation of the initial ~50% of the TBP into productive complexes. Although the second half of the TBP also rapidly forms non-productive complexes ($\leqslant$ one second), these incorrectly oriented I_1 complexes are readily reversible (dissociation $t_{1/2}$ = 28 seconds) and allow continued rapid sequential assembly of the PIC, with > 90% of the TBP proceeding along the initiation pathway within 80 seconds. Kinetic control thus allows transcription to be rapidly induced to the maximum rate of the promoter without the need to invoke interactions with TAF or activator proteins that prevent ternary complex formation in the incorrect orientation. In this model of transcription induction, crucial structural changes associated with formation of the final TBP-DNA complex would presumably accompany TFIIB binding, but would not be a prerequisite.

In conclusion, the fundamental principles of the initiation of gene transcription continue to be elucidated through the application of a wide range of experimental approaches. For a time-dependent process such as the sequential assembly of transcription pre-initiation complexes, an accurate description of the chemical processes that underlie the biology of these systems requires detailed, quantitative study of the reaction kinetics. This requirement is particularly true if the regulation of the process is important to understanding its biological function. What are the rate-determining steps in the reaction mechanism? Are there branch points and/or alternative pathways? The kinetics of transcription initiation is highly dependent on the interplay of the rate constants of the individual steps that comprise the reaction mechanism. The values of these constants can be obtained only from non-steady-state measurements. The study of the DNA sequence and TBP species-dependence of the kinetics of pre-initiation complex assembly is being pursued from this perspective.

Experimental procedures

Protein and DNA

Full-length *S. cerevisiae* TBP was prepared as described [Petri et al 1995 and Parkhurst et al 1996]. The two preparations of *S. cerevisiae* TBP used in this study were evaluated for consistency by both FRET and quantitative DNase I footprinting. Static and kinetic measurements made using the two preparations were identical within experimental error.

The double-labeled DNA containing the TATA sequence from the major late promo ter (TAMRA-$^{5'}$GGGC**TATAAAAG**GG$^{3'}$-fluorescein, or T*ML*F) is very similar to the probe used previously ([Parkhurst et al 1996], except that both the 3′ and 5′ linkages are six carbon atoms in length, in contrast to the previous use of a 4-carbon 3′ linkage and TAMRA replacing tetramethylrhodamine as the fluorescence acceptor. This oligomer was synthesized and both HPLC and PAGE-purified by Genosys Biotechnologies, Inc. (The Woodlands, TX). The complementary strand was synthesized in the University of Nebraska-Lincoln DNA synthesis facility. All studies were conducted in 10 mM Tris-HCl (p 7.4), 100 mM KCl, 2.5 Mm $MgCl_2$

and 1 mM $CaCl_2$ at the indicated temperature. Two preparations of the double-labeled DNA complexed with unlabeled complementary DNA (T*ML_{dpx}*F) demonstrated identical 44(±1.3)% decreases in the steady-state fluorescein emission intensity upon TBP binding (Figure 1).

Fluorescence stopped-flow measurements

Detailed considerations of Förster resonance energy transfer in general [Wu and Brand 1992, Wu and Brand 1994, Cheung 1991 and Lakowicz 1986] and as it pertains to this particular application [Parkhurst and Parkhurst 1994, Parkhurst and Parkhurst 1995a, Parkhurst and Parkhurst 1995b and Parkhurst et al 1996] have been published. The instrumentation and associated temperature control used for both the fluorescence stopped-flow and the steady-state fluorescence measurements were as described in the above cited references. Kinetic progress curves were obtained for DNA-TBP binding over a range of temperature and protein concentrations. Data were collected at five temperatures, 15, 20, 22.5, 25 and 30°C, using at least two TBP concentrations at each temperature. The DNA duplex solution was prepared using 40 nM T*ML*F and 60 nM complement, while the TBP solution was 400, 800 or 1600 nM (prior to stopped-flow mixing). The DNA and protein solutions were loaded into their respective stopped-flow drive syringes and temperature equilibrated; the baseline voltage of the photomultiplier output was recorded and data acquisition was initiated.

Preliminary experiments were conducted to determine the optimal time range for data collection; data points were subsequently obtained to include 10(±2) × $t_{1/2}$ of the slow phase of the reaction. Four or five independent kinetics progress curves were obtained at each combination of temperature and TBP concentration. Each set of curves was baseline corrected, averaged and smoothed, yielding a total of 11 averaged decay curves (five temperatures and two or three TBP concentrations at each temperature). The averaged curves were fit to mono-, bi- and triexponential decay models.

For any averaged curve, the error in the eigenvalue for the fast phase, λ_1, was determined from the variance in the values of λ_1 obtained for each of the four or five curves in a set. The fractional error in the amplitude of each averaged curve reflected an additional uncertainty, due to the slight variations in the amplitude assigned to photobleaching within a set of replicate curves. To circumvent this uncertainty, the amplitude errors were assumed to be proportional to the errors in the endpoints of the equilibrium titration curves, determined under the same experimental conditions, since the titration endpoints also corresponded to the overall change in the amplitude of the fluorescein emission associated with TBP binding. The average fractional error associated with the titration endpoints was determined using standard methods of error propagation.

Determination of equilibrium constants

Titrations of the T*ML_{dpx}*F with TBP were conducted at 15, 20, 22.5 and 25°C. The extent of binding was determined by monitoring changes in the steady-state fluorescence spectrum of 4 nM T*ML_{dpx}*F as a function of TBP concentration in six or seven increments from zero to 640 nM. The determination of the fraction of DNA bound, the equilibrium association constant (K_a) and the titration endpoint have been described [Parkhurst and Parkhurst 1995a and Parkhurst and Parkhurst 1994]. A van't Hoff plot was constructed using these values of K_a, and simple regression yielded the values of K_a used in the global analysis, including the value at 30°C obtained from extrapolation. This indirect determination of K_a at 30°C was deemed to be preferable to direct measurement due to the high affinity of TBP for the major late promoter at this temperature, resulting in a close correspondence of K_d and the concentration of T*ML_{dpx}*F. The error associated with each value of K_a was obtained using a standard expression for the estimated error in a value derived from a regression [Draper and Smith 1966].

Relaxation kinetic measurements

The value of the first-order relaxation constant, R, was determined at each of the five temperatures by adding a large excess of unlabeled 14-mer ML duplex (ML_{dpx}), hand-mixing and monitoring the increase in the steady-state emission of the T*ML_{dpx}*F as it was replaced by the ML_{dpx}. Numerical simulations were first conducted to determine the concentration of ML_{dpx} that would result in $\leqslant$ 5% of the T*ML_{dpx}*F bound at equilibrium, assuming identical values of K_a for TBP binding the labeled and unlabeled duplexes. The relaxation experiments were conducted at each temperature as follows: to 22 nM T*ML_{dpx}*F, TBP was added to a final concentration of 410–487 nM, depending upon the temperature, to ensure that $\geqslant$ 93% of the duplex was bound. ML_{dpx} was then added to a final concentration of 8.2 μM. A total of 25–30 emission spectra were collected from 500–600 nm as a function of time for 50 minutes beginning ca seven seconds following the addition and mixing. The fraction of bound T*ML_{dpx}*F was determined, as described [Parkhurst and Parkhurst 1995a], for each time point at a given temperature. These data were fit to mono- and biexponential decay models.

Global analysis

The collective kinetic and thermodynamic data were globally fit by a general minimization program, using a subroutine that integrates the differential equations appropriate to any given mechanism. The data used in the global analyses included the 11 stopped-flow kinetic curves and the values of K_a and R determined at the five temperatures. The correspondence of the data to one and two-intermediate models ((2) and (3), respectively) was explored. Details of fitting to the two-intermediate model are presented first; differences pertaining to fitting to the one-intermediate model follow.

For the two-intermediate model, values were determined for the six rate constants, k_1–k_6, at 30°C, as well as the six activation enthalpies associated with those rate constants. From the optimal values for these fitted parameters, the values of the rate constants at each of the other four temperatures as well as $\Delta S^{\circ\neq}$ and $\Delta G^{\circ\neq}$ for each rate constant were calculated according to standard transition state theory. Three quantum yield ratios were also determined in the global fit, which relate the fluorescein emission intensity for T*ML_{dpx}*F in I_1, I_2 and the final complex to that of the unbound probe.

In the first segment of the global analysis, each of the 11 measured stopped-flow binding curves was constructed using the amplitudes (α_1 and α_2) and eigenvalues (λ_1 and λ_2) that described each averaged curve. Each curve was represented by 41 points, which were evenly distributed between the slower and faster phases. The first 21

points (including that for time zero) extended in time to $5 \times t_{1/2}$ for λ_1, and the second 20 points, corresponding to λ_2, extended from the end of the fast phase to 97% completion, which was determined using a Newton-Raphson procedure. The error in these points was determined from smoothed replicate runs, with an average error of 0.023. To obtain the calculated values for the individual points describing each decay curve, the rate constant matrix for the two-intermediate model was first constructed. A quartic root finder was used to solve at each iteration for the values of λ_1, λ_2 and $\lambda_3(\lambda_4 = 0)$, and the pre-exponential terms calculated. The species concentrations could then be determined as a function of time, and, finally, the response function, $R(t)$, was constructed, having the general form $R(t) = [F(t) - F(\infty)]/F(0)$, where $F(t)$, $F(\infty)$ and $F(0)$ are the fluorescence emission of the donor fluorescein at any time, t, at equilibrium, and at time zero, respectively. Within the theoretical response function, each species concentration was multiplied by the relative quantum yield of that species, as described above, with the quantum yield of the fluorescein in the unbound $T*ML_{dpx}*F$ defined as 1. The quality of the fit to the 11 averaged stopped-flow decay curves, σ_{curve}^2, was evaluated as follows:

$$\sigma^2_{\text{curve}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m}\frac{(R^0_{ji} - R^C_{ji})^2}{\sigma^2_{\text{curve},\,j}}\right) \tag{7}$$

where n is the total number of curves; m is the total number of points per curve; R^o and R^c are the values of the observed and calculated response functions, respectively; and i and j index the curves and the points per curve, respectively. The calculated values of λ_1 that fell outside the 95% confidence region for the observed values contributed to a weighted sum of squared residuals (σ_λ^2), which was incorporated into the overall sum of squared residuals for the global analysis (σ_{global}). In an entirely analogous manner, the total change in amplitude, $\alpha_1+\alpha_2$, for each curve was evaluated, and the associated weighted sum of squared residuals (σ_{amp}^2) contributed to σ_{global}.

In the next segment of the global analysis, but within the same iterative step, the calculated values of K_a and R at each temperature were evaluated. Calculated values of K_a outside of the 95% confidence region, determined using the F statistic, contributed to σ_{global} *via* their weighted sum (σ_K^2). When the experimental values obtained for the relaxation constant, R, at all five temperatures were subjected to the statistical F-test at the 95% confidence limit, the null hypothesis obtained; within error, all R values were the same. Calculated values of R falling outside the two extremes of the measured values (including the respective errors of those values) contributed to σ_{global} *via* the weighted sum of squared residuals for R (σ_R^2). The overall quality of the global fit, σ_{global}^2, was determined using the weighted residuals from five sources: the points describing each of the 11 stopped-flow curves, the first eigenvalue associated with each averaged curve, the amplitude change of each curve, the five equilibrium constants and the five relaxation constants (equation (1).

Global fitting of the stopped-flow, equilibrium and relaxation data to the one-intermediate model (equation (2) was conducted in a manner entirely analogous to that for the two-intermediate model. Values were determined for each of the four rate constants, k_1–k_4, at 30°C, the four associated activation enthalpies, as well as the relative quantum yields of the intermediate and final complexes. The one-intermediate mechanism was integrated in closed form to yield expressions for λ_1 and $\lambda_2(\lambda_3 = 0)$ and the coefficients and values were calculated for the points describing each stopped-flow decay curve. The value of σ_{global} was determined exactly as described for the two-intermediate model.

Error estimates for parameters determined in global analysis

Errors for each parameter were estimated by identifying the parameter sets encompassed within the 68% confidence region for σ_{global}. First, the value of each single parameter was sequentially changed and σ_{global} optimized at each step. The parameter values corresponding to the 68% confidence region for σ_{global} gave the upper and lower error bounds for that parameter. These univariate limits together defined the 15-dimensional rectangular prism corresponding to the 68% confidence region. For the 15-dimensional hyper-ellipsoid corresponding to the joint 68% confidence region within the rectangular prism, nearly all of the volume is within a very narrow boundary of the surface. A Golden Section search [Walsh 1975] was used to identify this boundary. In this procedure, a random point within the rectangular prism is chosen, and a search in which all parameters are varied is conducted along a trajectory from the value of the global minimum (the optimal σ_{global}) toward the randomly chosen point, until the 68% confidence limit of σ_{global} intersects the hyper-ellipsoid. The 15 parameter values associated with that intersection represent limits on the 15 respective optimal values. One thousand such intersections were randomly identified, yielding 1000 independent sets of 15 parameter values that coincided with the joint σ_{global}68% confidence region. Among the 1000 values thus determined for each parameter, very few outlying points were obtained. To minimize their effect, the error estimate for the standard deviation of each parameter was defined as the average of the extreme upper and lower 5% of the 1000 values. (A Monte Carlo search was unsuccessful in determining parameter values within the joint 68% hyper-ellipsoid region because the volume of that region was $<10^{-6}$ of the enclosing rectangular prism.)

To determine whether these error estimates accurately reflected the errors in the parameters derived from the global fit, the observed errors in the parameter k_1 were determined directly from the stopped-flow data and compared with those obtained from the error analysis just described. For certain conditions, the fast eigenvalue $\lambda_1 \approx k_1[\text{TBP}]+k_2$, so that for a small value of k_2, $k_1 \approx \lambda_1/[\text{TBP}]$. A few such data sets, each at a single temperature and [TBP], were identified for which these two approximations held; the fractional variance of λ_1 within a data set gave a reliable indication of the upper and lower error bounds for the associated k_1. The limits determined for $k_1$from the global error analysis described above corresponded very well with the simple error estimates made using the experimental data. It should be noted that global analysis provides improvement in error estimates relative to those obtained from non-global analyses, so that the errors reported herein should not be underestimated.

Standard variance-covariance matrix analysis proved inappropriate for this highly non-linear problem where the minimum value for a given parameter was not symmetrically situated with regard to the hyper-ellipsoid defining the 68% confidence boundary, resulting in error estimates that were too large. Error estimation was also ex-

plored by initiating the analysis from the global minimum and varying a single parameter, holding all other parameters constant until the 68% σ_{global} confidence boundary of the 15-dimensional rectangular prism was reached. The resulting errors were underestimated according to the errors estimated for k_1 as just described.

DNA binding by TBP at stoichiometric DNA concentrations

A series of kinetics experiments were conducted at stoichiometric concentrations of DNA and TBP. In these experiments, solutions were prepared containing 0.040, 0.5, 1.5, 4 and 8 μM T*ML_{dpx}*F at 30°C. For each solution, an emission spectrum was obtained for the labeled duplex. TBP was then diluted directly from cold 43.9 μM stock to a final concentration equal to that of the DNA, the solution was hand-mixed, and a second emission spectrum collected within 20 seconds of TBP addition. At the high concentrations of duplex, the T*ML_{dpx}*F was present as a trace probe, comprising 0.5–1.0% of the total duplex, together with unlabeled ML_{dpx} sufficient to reach the desired total concentration. To confirm that identical results were obtained with the T*ML_{dpx}*F used as a trace probe, the following procedure was done. An emission spectrum was collected for a solution containing 40 nM T*ML_{dpx}*F. TBP was added to a final concentration of 40 nM, and a second spectrum was taken following completion of binding. This procedure was repeated using 4 nM T*ML_{dpx}*F+36 nM unlabeled duplex (10% double-labeled probe) and also 2 nM T*ML_{dpx}*F+38 nM unlabeled duplex (5% probe). [Parkhurst et al 1994]

Acknowledgements

We gratefully acknowledge support for this work from the Center for Biotechnology and the College of Arts and Sciences, University of Nebraska-Lincoln (L.J.P.), the Nebraska State Health Department 98-34 (L.J.P.), a Fellowship from the Program in Mathematics and Molecular Biology at Florida State University supported by National Science Foundation grant DMS-9406348 (R.M.R.), and National Institutes of Health grants GM39929 and GM51506 (M.B). M.B. thanks Ian Willis and Robyn Moir for many illuminating discussions.

References

Burley , S. K., and R. G. Roeder, Biochemistry and structural biology of transcription factor IID (TFIID). *Annu. Rev. Biochem.* **65** (1996), pp. 760–799.

Cantor, C. R., and P. R. Schimmel. In: *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules*, W. H. Freeman and Company, New York (1980), pp. 284–285.

Cheung, H. C., Resonance energy transfer. In: J. R. Lakowicz, Editor, *Topics in Fluorescence Spectroscopy* **vol. 2**, Plenum Press, New York (1991), pp. 157–171.

Coleman, R. A., and B. F. Pugh, Slow dimer dissociation of the TATA binding protein dictates the kinetics of DNA binding. *Proc. Natl Acad. Sci. USA* **94** (1997), pp. 7221–7226.

Coleman, R. A., A. K. Taggart, L. R. Benjamin, and B. F. Pugh, Dimerization of the TATA binding protein. *J. Biol. Chem.* **270** (1995), pp. 13842–13849.

Cox, J. M., M. M. Hayward, J. F. Sanchez, L. D. Gegnas, S. vander Zee, J. H. Dennis, P.B. Sigler, and A. Schepartz, Bidirectional binding of the TATA box binding protein to the TATA box. *Proc. Natl Acad. Sci. USA* **94** (1997), pp. 13475–13480.

Daugherty, M. A., M. Brenowitz, and M. G. Fried, The TATA-binding protein from *Saccharomyces cerevisiae* oligomerizes in solution at micromolar concentrations to form tetramers and octamers. *J. Mol. Biol.* **285** (1999), pp. 1389–1399.

Draper, N. R., and H. Smith. In: *Applied Regression Analysis*, John Wiley & Sons, Inc, New York (1966), pp. 21–24.

Gurney, R. W. In: *Ionic Processes in Solution*, McGraw Hill, New York (1953), pp. 89–91.

Hernandez, N., TBP, a universal eukaryotic transcription factor?. *Genes Dev.* **7** (1993), pp. 1291–1308.

Hoopes, B. C., J. P. LeBlanc, and D. K. Hawley, Kinetic analysis of yeast TFIID-TATA box complex formation suggests a multi-step pathway. *J. Biol. Chem.* **267** (1992), pp. 11539–11547.

Hoopes, B. C., J. F. LeBlanc, and D. K. Hawley, Contributions of the TATA box sequence to rate-limiting steps in transcription initiation by RNA polymerase II. *J. Mol. Biol.* **277** (1998), pp. 1015–1031.

Horikoshi, M., C. Bertuccioli, R. Takada, J. Wang, T. Yamamoto, and R.G. Roeder, Transcription factor TFIID induces DNA bending upon binding to the TATA element. *Proc. Natl Acad. Sci. USA* **89** (1992), pp. 1060–1064.

Juo, Z. S., T. K. Chiu, P. M. Leibermann, I. Baikalov, A. J. Berk, and R. E. Dickerson, How proteins recognize the TATA box. *J. Mol. Biol.* **261** (1996), pp. 239–254.

Kim, J. L., and S. K. Burley, 1.9 Å resolution refined structure of TBP recognizing the minor groove of TATAAAAG. *Nature Struct. Biol.* **1** (1994), pp. 638–653.

Kim, J. L., D. B. Nikolov, and S.K. Burley, Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365** (1993), pp. 520–527.

Kim, Y., J. H. Geiger, S. Hahn, and P. B. Sigler, Crystal structure of a yeast TBP/TATA-box complex. *Nature* **365** (1993), pp. 512–520.

Knutson, J. R., J. Beechem, and L. Brand, Simultaneous analysis of multiple fluorescence decay curves: a global approach. *Chem. Phys. Letters* **102** (1983), pp. 501–507.

Lakowicz, J.R. In: *Principles of Fluorescence Spectroscopy*, Plenum Press, New York (1986), pp. 305–337.

Lee , T. I., and R. A. Young, Regulation of gene expression by TBP-associated proteins. *Genes Dev.* **12** (1998), pp. 1398–1408.

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

McKnight, S., Transcription revisited: a commentary on the 1995 Cold Spring Harbor Laboratory meeting, "mechanisms of eukaryotic transcription". *Genes Dev.* **10** (1996), pp. 367–381.

Nikolov, D. B., H. Chen, E. D. Halay, A. Hoffman, R. G. Roeder, and S.K. Burley, Crystal structure of a human TATA box-bind-

ing protein/TATA element complex. *Proc. Natl Acad. Sci. USA* **93** (1996), pp. 4862–4867.

Parkhurst, L. J., and K. M. Parkhurst, Changes in the end-to-end distance distribution in an oligonucleotide following hybridization. *SPIE* **2137** (1994), pp. 475–483.

Parkhurst, K. M., and L. J. Parkhurst, Kinetic studies by fluorescence resonance energy transfer employing a double-labeled oligonucleotide: hybridization to the oligonucleotide complement and to single-stranded DNA. *Biochemistry* **34** (1995), pp. 285–292.

Parkhurst, K. M., and L. J. Parkhurst, Donor-acceptor distance distributions in a double-labeled fluorescent oligonucleotide both as a single stand and in duplexes. *Biochemistry* **34** (1995), pp. 293–300.

Parkhurst, K. M., R.E. Hileman, D. Saha, N. K. Gupta, and L. J. Parkhurst, Thermodynamic characterization of the cooperativity of 40S complex formation during the initiation of eukaryotic protein synthesis. *Biochemistry* **33** (1994), pp. 15168–15177.

Parkhurst, K., M. Brenowitz, and L. J. Parkhurst, Simultaneous binding and bending of promoter DNA by TBP: real time kinetic measurements. *Biochemistry* **35** (1996), pp. 7459–7465.

Perez-Howard, G., T. Weil, and J. M. Beechem, Yeast TATA binding protein interaction with DNA: fluorescence determination of oligomeric state, equilibrium binding, on-rate, and dissociation kinetics. *Biochemistry* **34** (1995), pp. 8005–8017.

Petri, V., M. Hsieh, and M. Brenowitz, Thermodynamic and kinetic characterization of the binding of the TATA binding protein to the adenovirus E4 promoter. *Biochemistry* **34** (1995), pp. 9977–9984.

Petri, V., M. Hsieh, E. Jamison, and M. Brenowitz, DNA sequence-specific recognition by the 'TATA' binding protein: promoter dependent differences in the thermodynamics and kinetics. *Biochemistry* **37** (1998), pp. 15842–15849.

Sethy-Coraci, I., R. D. Moir, A. López-de-León, and I. M. Willis, A differential response of wild type and mutant promoters to $TFIIIB_{70}$ over expression *in vivo* and *in vitro*. *Nucl. Acids Res.* **26** (1998), pp. 2344–2352.

Starr, D. B., B. C. Hoopes, and D. K. Hawley, DNA bending is an important component of site-specific recognition by the TATA binding protein. *J. Mol. Biol.* **250** (1995), pp. 434–446.

Walsh, G. R. In: *Methods of Optimization*, John Wiley & Sons, New York (1975), pp. 91–93.

Werner, M. H., A. M. Grononborn and G. M. Clore, Intercalation, DNA kinking, and the control of transcription. *Science* **271** (1996), pp. 778–784.

Wu, P., and L. Brand, Orientation factor in steady-state and time resolved resonance energy transfer measurements. *Biochemistry* **31** (1992), pp. 7939–7947.

Wu, P., and L. Brand, Resonance energy transfer: methods and applications. *Anal. Biochem.* **218** (1994), pp. 1–13.