Kinetic and Thermodynamic Basis of Promoter Strength: Multiple Steps of Transcription Initiation by T7 RNA Polymerase Are Modulated by the Promoter Sequence†

Rajiv P. Bandwar, Yiping Jia, Natalie M. Stano, and Smita S. Patel*

Department of Biochemistry, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854

Received October 10, 2001; Revised Manuscript Received January 18, 2002

ABSTRACT: Transcription initiation by T7 RNA polymerase (T7 RNAP) is regulated by the specific promoter DNA sequence that is classically divided into two major domains, the binding domain (−17 to −5) and the initiation domain (−4 to +6). The occurrence of nonconsensus bases within these domains is responsible for the diversity of promoter strength, the basis of which was investigated by studying T7 promoters with changes in the promoter specificity region (−13 to −6) of the binding domain and/or the melting region (−4 to −1) of the initiation domain. The transient state kinetics and thermodynamic studies revealed that multiple steps in the pathway of transcription initiation are modulated by the promoter DNA sequence. Three base changes in the promoter specificity region at −11, −12, and −13, found in the natural ϕ3.8 promoter, reduced the overall affinity of the T7 RNAP for the promoter DNA by 2−3-fold and decreased the rate of pppGpG synthesis, the first RNA product. Promoter opening is thermodynamically driven in T7 RNAP, and a single base change in the melting region (TATA to TAAA) decreased the extent of open complex generated at equilibrium. This base change in the melting region also increased the K_d of (+1) GTP and the dissociation rate of pppGpG. Thus, transcription initiation at various T7 promoters is differentially regulated by initiating GTP concentration. The specificity and melting regions of T7 promoter DNA act both independently and synergistically to affect distinct steps of transcription initiation. Although each step in the initiation pathway is affected to a small degree by promoter sequence variations, the cumulative effect dictates the overall promoter strength.

The promoter DNA sequence contains all the information to direct and regulate the synthesis of RNA during transcription. The promoters of bacteria or bacteriophages such as T7, T3, and SP6 contain a consensus sequence that is recognized by the respective RNAP.[1] Distinct domains of the promoter DNA are recognized at different stages of initiation, and when a functional initial complex, referred to as the closed complex, is formed, it undergoes multiple conformational changes during which a DNA region close to the initiation site is melted. Kinetic and thermodynamic studies show that transcription initiation is a multistep process in both *Escherichia coli* and T7 RNAP (*1*−*3*). To understand how initiation is regulated, one needs to identify the elementary steps, determine their rate constants, and determine how these steps and their rate constants are affected by promoter sequence and/or by transcription factors. T7 RNAP, being a relatively simple enzyme, is an ideal system to study and to derive principles of how transcription initiation can be regulated by the promoter DNA sequence.

The crystal structures of the T7 RNAP−DNA complex reveal a detailed view of the promoter in the open complex (*4*−*7*). In these structures, the promoter binding domain (−17 to −5) is duplex and the melting region (−4 to −1) is single stranded and suspended toward the active site pocket. Mutagenesis and crystal structural studies indicate that the T7 RNAP−DNA complex is stabilized by specific interactions of the protein with bases in the specificity region as well as the melting region of the promoter (*8*−*10*). These interactions confer specificity and somehow dictate the transcriptional efficiency of the T7 RNAP, both of which can be altered by base changes in the promoter DNA (*8*, *11*−*13*). The promoters of bacteriophage T7 share a consensus sequence of 23 base pairs from −17 to +6 relative to the initiation site +1 (*14*). The class III promoters possess the exact consensus sequence and are considered strong promoters whereas the class II promoters have divergent bases and are weak promoters (*15*−*18*). It is well-known that the promoter sequence itself is responsible for modulating transcription by T7 RNAP along with the negative regulator, T7 lysozyme protein (*19*−*23*). A number of studies reveal the hierarchy of base pair preference in T7 promoters, and measurement of the overall efficiency of transcription reveals the importance of a particular base in the promoter sequence (*24*, *25*). Although it is obvious that promoter sequence is responsible for promoter strength, the molecular basis for the promoter strength is not fully understood (*1*, *26*, *27*). The promoter sequence can dictate the efficiency of initiation and promoter clearance and control the synthesis of the final RNA transcript. Obvious targets of regulation

† This research was supported by NIH Grant GM51966.

* Corresponding author. Telephone: 732-235-3372. Fax: 732-235-4783. E-mail: patelss@umdnj.edu.

[1] Abbreviations: RNAP, RNA polymerase; 2-AP, 2-aminopurine; bp, base pair; t, template; nt, nontemplate; ds, double stranded; p-ds, partially double stranded.

10.1021/bi0158472 CCC: $22.00
Published on Web 02/20/2002

are the rate-limiting steps and those with an unfavorable equilibrium constant. Studies have not been done to systematically analyze the effect of promoter sequence variation on each step of initiation.

We have focused our studies on transcription initiation to understand promoter strength using T7 RNAP as a model system. Considering that bacteriophage RNAPs show similarity to mitochondrial and chloroplast RNAPs (*28*), detailed studies of T7 RNAP can be valuable in understanding the mechanism and regulation of transcription in higher organisms. We have developed fluorometric and radiometric methods to measure the transient state kinetics of each step during transcription initiation (*2*, *3*, *29*). We apply these methods to examine the basis of promoter strength differences between ϕ3.8, a weaker promoter, and ϕ10, a strong promoter, and a modified ϕ3.8 promoter. The results of these studies indicate that regulation of transcription occurs at several steps during initiation, and both intrinsic rate constants and equilibrium constants are modulated by the promoter sequence. Some of the steps that are regulated include the equilibrium constant of closed and open complexes, extent of promoter melting, the K_d of initiating NTP, and the rate of the first phosphodiester bond formation reaction.

MATERIALS AND METHODS

Protein. T7 RNAP was purified using a modified procedure from *E. coli* strain BL21/pAR1219 (*30*). The enzyme was >95% pure and free of contaminating exonuclease activity. The protein was stored at −80 °C after dialyzing against the buffer containing 50% glycerol (v/v), 20 mM sodium phosphate, pH 7.7, 1 mM Na_3EDTA, 1 mM dithiothreitol, and 100 mM NaCl. The enzyme concentration was determined from its absorbance and molar extinction coefficient of 1.4×10^5 M^{-1} cm^{-1} at 280 nm.

Promoter DNA. The 2-AP-modified and unmodified oligodeoxynucleotides were purchased from Integrated DNA Technologies (Coralville, IA). The chemically synthesized DNAs were purified by gel electrophoresis. The DNA band corresponding to the desired length was excised, electroeluted, ethanol precipitated, resuspended in water, and stored at −20 °C. After purification, the ssDNAs of the desired length were the only band detectable, and we estimate that the purity of ssDNAs was >95%. The DNA concentrations were determined from the calculated molar extinction coefficients, as described previously (*29*, *31*). The dsDNAs were prepared by annealing the individual ssDNA strands. The exact ratio of the two ssDNAs used to prepare the dsDNA was determined from titration experiments performed on a 16% native polyacrylamide gel that resolves the dsDNA from the ssDNAs.

Fluorescence Measurements. The equilibrium DNA binding experiments were carried out at 25 °C in a 3 mL quartz cuvette on a FluoroMax-2 spectrofluorometer (Jobin Yvon-Spex Instruments S.A., Inc.) using the DataMax software program (*31*). Fluorescence titrations were carried out by adding small aliquots (2−4 μL) of t(−4) 2-AP DNA to a constant amount of T7 RNAP (0.5 μM) in 2.5 mL of buffer A (50 mM Tris−acetate, 50 mM sodium acetate, 10 mM magnesium acetate, 5 mM dithiothreitol). The corrected fluorescence (F_c) was obtained after correction for inner filter and volume using the equation:

$$F_c = F_{obs}\left(\frac{\upsilon_f}{\upsilon_0}\right) \times 10^{0.5(Abs_{ex}+Abs_{em})} \quad (1)$$

where F_{obs} is the observed fluorescence intensity, υ_f is the final volume of the solution, υ_0 is the initial volume, Abs_{ex} is the absorbance of the T7 RNAP−DNA solution at the 2-AP excitation wavelength of 315 nm, and Abs_{em} is the absorbance of the same solution at the 2-AP emission wavelength of 370 nm. The corrected fluorescence was plotted as a function of total DNA concentration, and the data were fit to the quadratic equation (eq 2) to obtain the K_d values:

$$F_c = F_{max}\left(\frac{(K_d + E_t + D_t) - \sqrt{(K_d + E_t + D_t)^2 - 4(E_tD_t)}}{2}\right) \quad (2)$$

where E_t and D_t are the concentrations of total enzyme and total DNA, K_d is the dissociation constant of the T7 RNAP−DNA complex, and F_{max} is the maximum fluorescence change at saturating [DNA].

To measure the extent of open complex formed at equilibrium, the fluorescence intensity of buffer A, T7 RNAP (4 μM), DNA (1 μM), or the T7 RNAP−DNA mixture (4 μM T7 RNAP + 1 μM DNA) was measured after excitation at 315 nm and emission at 370 nm. The fluorescence intensity of the T7 RNAP−DNA (ED) complex was determined after correcting the fluorescence of free DNA and T7 RNAP using the formula fluorescence = F_c(f) − F_c(nf), where F_c(f) and F_c(nf) are the fluorescence intensities of the ED complex with fluorescent (t(−4) 2-AP modified) and nonfluorescent (unmodified) DNA, respectively.

Stopped-Flow Kinetics. The stopped-flow experiments were carried out using a SF-2001 spectrophotometer equipped with a photomultiplier detection system from KinTek Corp. (Austin, TX). Varying concentrations of T7 RNAP (40 μL) from one syringe were rapidly mixed with a constant amount (0.15 μM final) of t(−4) 2-AP DNA (40 μL) from a second syringe at 25 °C at a flow rate of 6.0 mL/s. The sample was excited with light at 315 nm wavelength, and the progress of the reaction was monitored by measuring the intensity of the emission using a cut-on filter >360 nm (WG360). About 5−10 traces (one trace per mixing) were averaged at each concentration of T7 RNAP and the time courses at all concentrations best fit to a single exponential equation, $F = A(1 - e^{-k_{obs}t}) + C$, where F is the 2-AP fluorescence at time t, A is the amplitude of the fluorescence change, k_{obs} is the observed rate constant, and C is the fluorescence at $t = 0$. The k_{obs} thus obtained under pseudo-first-order conditions was plotted as a function of total T7 RNAP concentration, $[E]_t$, and fit to eq 3 using SigmaPlot (Jandel Scientific) to obtain the maximum observed rate (k_{max}), the $K_{1/2}$, and k_{off} from the y-intercept:

$$k_{obs} = \frac{k_{max}[E]_t}{K_{1/2} + [E]_t} + k_{off} \quad (3)$$

The K_d for GTP was determined by measuring the kinetics of fluorescence change after mixing a complex of T7 RNAP

(0.45 μM final) and nt(+4) 2-AP DNA (0.15 μM final) with varying concentrations of GTP (or GTP + GMP) in a stopped-flow instrument, as described above. The observed rates were plotted as a function of final [GTP], and the dependency was fit to the hyperbolic equation:

$$k_{obs} = \frac{V_{max}[GTP]}{K_d + [GTP]} \quad (4)$$

where V_{max} is the observed rate of a conformational change upon GTP binding and K_d is the equilibrium dissociation constant of GTP.

Pre-Steady-State Kinetics of RNA Synthesis. The pre-steady-state kinetic experiments were carried out on a rapid chemical quench-flow instrument (KinTek Corp., Austin, TX) (*2*). A typical quench-flow experiment was carried out as follows. T7 RNAP (15 μM final) and promoter DNA (10 μM final) were preincubated in buffer A in one syringe. In another syringe was loaded the substrate solution containing GTP and [γ-^{32}P]GTP (from Amersham Life Science). The temperature was maintained constant at 25 °C using a water bath. The reaction was initiated by rapidly mixing equal volumes of the two solutions in the quench-flow instrument. After predetermined time intervals, the reactions were quenched by rapidly mixing with 1 N HCl from a third syringe. Chloroform was then added, and the reactions were neutralized using an appropriate volume of 0.25 M Tris base and 1 M NaOH within 1 min. The RNA products were resolved by electrophoresis at 55 °C (110 W) on a highly cross-linked 23% polyacrylamide/3 M urea gel on a Bio-Rad sequencing gel apparatus (0.25 mm thick spacers and comb). The gel was exposed to a phosphor screen for about 15 h and scanned on a PhosphorImager instrument (Molecular Dynamics), and the RNA products and the substrate were quantitated using the ImageQuaNT program.

The pre-steady-state kinetic data were fit to the burst equation (eq 5) using SigmaPlot software:

$$y = A(1 - e^{-kt}) + bt + C \quad (5)$$

where A is the burst amplitude, k is the exponential burst rate constant, b is the linear steady-state rate constant, and C is the y-intercept. The GTP concentration dependence of the burst rate was fit to the hyperbolic equation (eq 4).

RESULTS

The ϕ3.8 and ϕ10 promoter sequence from −21 to +19 was incorporated into 40 bp oligodeoxynucleotide constructs as shown in Figure 1. These two natural T7 promoters were selected as representatives of class II and class III promoters. The two major "domains", binding (−17 to −5) and initiation (−4 to +6), have been further subdivided into specificity (−13 to −6) and melting (−4 to −1) "regions". To explore distinct effects of the promoter specificity and melting regions on steps of initiation, a variant (A−2T) of ϕ3.8 promoter was also prepared. The ϕ3.8(A−2T) promoter has the same sequence as ϕ3.8 in the binding domain, only the −2A base in the initiation domain was changed to −2T to restore the consensus sequence TATA. To enable measurement of the kinetics and thermodynamics of promoter opening and GTP binding, the t(−4) and nt(+4) adenines were individually substituted with the fluorescent base 2-AP.

ϕ10
AAATTAAT|**ACG**|ACTCAC|**TATA**|G̲GGAGACCACAACGGTTTC
TTTAATTA|TGC|TGAGTG|ATAT|C̄CCTCTGGTGTTGCCAAAG

ϕ3.8
TAATTAAT|**tga**|ACTCAC|**TAaA**|G̲GGAGACCACAGCGGTTTC
ATTAATTA|act|TGAGTG|ATtT|C̄CCTCTGGTGTCGCCAAAG

ϕ3.8(A-2T)
TAATTAAT|**tga**|ACTCAC|**TATA**|G̲GGAGACCACAGCGGTTTC
ATTAATTA|act|TGAGTG|ATAT|C̄CCTCTGGTGTCGCCAAAG

FIGURE 1: Sequences of synthetic T7 DNA promoters. The upper row of each 40 bp double-stranded DNA represents the 5′−3′ sequence of the nontemplate strand of the promoter DNA. The sequences shown are from −21 to +19 relative to the start site, G, at +1. The promoter specificity region (−13 to −11) and melting region (−4 to −1) sequences for promoter strength comparison are shown in bold letters. The base pair changes in ϕ3.8 and ϕ3.8(A−2T) promoters from the consensus sequence (−17 to +6) of ϕ10 promoter are shown in lowercase letters.

We have shown previously that even multiple 2-AP substitution (from −4 to +4) has no deleterious effect on transcription (*29*).

Nonconsensus Promoters Bind T7 RNAP with a Weaker Affinity Than Consensus Promoters. The affinity of T7 RNAP for the promoter DNA has been estimated by a number of methods, including fluorophore-labeled DNAs measuring either the anisotropy changes or fluorescence intensity changes of the DNA (*29*, *32*). Here we have used the fluorescence method to determine the relative binding affinities of the three promoters for T7 RNAP. To measure the K_d of promoter DNA by the fluorescence method, the promoter DNA was modified with the fluorescent adenine analogue, the 2-AP. The fluorescence of 2-AP base (incorporated in the TATA sequence, −4 to −1) increases when T7 RNAP binds the promoter DNA (*29*, *31*, *33*). Previously, we measured the K_d's of ϕ10 and ϕ3.8 promoters using multiple 2-AP's incorporated in the −4 to +4 region (*29*); the ϕ3.8 promoter exhibited 2-fold weaker K_d relative to the ϕ10 promoter. Recently, we reported a peculiarly large fluorescence increase from a single 2-AP incorporated at the t(−4) position (*31*) and attributed to 2-AP unstacking from the neighboring t(−5) guanine when the TATA sequence melts during open complex formation (*7*). This fluorescence increase provides a sensitive signal for measuring the affinity of the promoter DNA for the T7 RNAP (and also the kinetics of initiation, described below) with minimal perturbation. In the fluorometric titration, a constant amount of T7 RNAP was titrated with increasing concentrations of t(−4) 2-AP DNA, and the fluorescence of 2-AP was measured at equilibrium. The corrected fluorescence was plotted versus [2-AP DNA], and the equilibrium DNA binding isotherms (Figure 2) were fit to a single-site DNA binding model (eq 2) to obtain the overall K_d for each promoter. Both ϕ3.8 and ϕ3.8(A−2T) promoters show a 2-fold weaker affinity for T7 RNAP, with K_d close to 1 μM, relative to ϕ10 promoter with a K_d close to 0.5 μM. The maximum fluorescence increase (F_{max}) is different for each promoter, with ϕ10 showing the largest increase and ϕ3.8 the lowest.

The measured dissociation constants are macroscopic parameters that provide the overall K_d of the complexes generated when T7 RNAP binds to the promoter DNA. We have proposed that the pathway of open complex formation

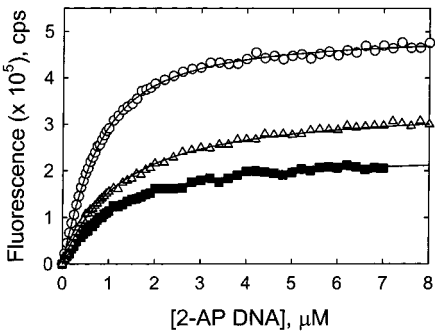


FIGURE 2: Equilibrium binding of ϕ10, ϕ3.8, and ϕ3.8(A−2T) DNA promoters. The fluorescence of the T7 RNAP−DNA complex using 2-AP DNA was measured at increasing DNA concentration and at constant T7 RNAP concentration (0.5 μM) at 25 °C for ϕ10 (○), ϕ3.8 (■), and ϕ3.8(A−2T) (△). The corrected fluorescence versus 2-AP DNA concentration was fit to eq 2 to obtain the K_d (± standard error) of ϕ10 (0.47 ± 0.02 μM), ϕ3.8 (0.95 ± 0.10 μM), and ϕ3.8(A−2T) (1.13 ± 0.08 μM).

goes through at least one intermediate, a closed complex ED_c that is formed before the formation of the open complex, ED_o (*31*). The measured K_d values are therefore close to $K_1(1 + K_2)$, where K_1 is the association constant of the E + D ⇔ EDc reaction and K_2 is the association constant of the $ED_c \Leftrightarrow ED_o$ reaction. The higher K_d values of ϕ3.8 promoters indicate that either the association constant K_1 or both K_1 and K_2 are affected by the promoter sequence. The K_d's of p-ds ϕ10 and p-ds ϕ3.8 promoters were also estimated from k_{off} and k_{on} values obtained from pre-steady-state stopped-flow kinetics reported earlier (*29*). The p-dsDNAs contain only the ds binding domain and lack the melting and coding regions and therefore provide an intrinsic K_d of the promoter region. The binding affinity for p-ds ϕ3.8 was found to be 60-fold weaker than p-ds ϕ10 (18 vs 0.3 nM). The weaker affinity of p-ds ϕ3.8 was reflected in a 70-fold faster dissociation rate (k_{off}) than p-ds ϕ10 promoter due to the nonconsensus promoter binding sequence only.

The Extent or Amount of Open Complex Is Lower with the Nonconsensus Promoters. Promoter melting is a critical step that can be controlled at the kinetic or the thermodynamic level by RNAP−DNA interactions that occur during closed complex formation. The preinitiation open complex formation has been investigated in T7 RNAP using various methods such as measurement of linking numbers in a plasmid relaxation assay (*34*), $KMnO_4$ oxidation of unpaired thymines in the TATA region (*35*, *36*), and measurement of 2-AP fluorescence in DNA (*31*). All methods indicate that although most of the T7 RNAP is bound to DNA, only part of the T7 RNAP−DNA complex is in the open form, as ED_o. In addition, it was observed that when GTP, the initiating nucleotide, and ATP were added, the open complex was readily detected by increase in $KMnO_4$ modification of the TATA region (*36*), change in the plasmid linking number (*34*), and as an increase in the fluorescence of 2-AP DNA in the TATA region (Stano and Patel, unpublished results). We interpret this result to mean that, in the absence of initiating nucleotide, most of the T7 RNAP−DNA complex is present as ED_c that contains DNA in the closed or unmelted form, which is in equilibrium with ED_o, and GTP + ATP binding drives ED_c to ED_o conversion. The promoter sequence could affect both the equilibrium and kinetics of preinitiation open complex formation, which we have measured here for all three promoters using the 2-AP fluorescence assay. The 2-AP fluorescence has been a subject of numerous theoretical and experimental studies (*37*, *38*), especially because 2-AP has proven to be a valuable probe in studying the dynamics of nucleic acids and nucleic acid binding proteins. These studies of 2-AP fluorescence indicate that quenching of 2-AP fluorescence in DNA is mainly due to base stacking interactions and collisions with neighboring bases, and the fluorescence of 2-AP in DNA is relatively insensitive to base-pairing or H-bond interactions. Thus, the fluorescence intensity changes of 2-AP in promoter DNA can be interpreted in terms of the degree of 2-AP base unstacking during open complex formation. The fluorescence method is quantitative and provides sensitive measurement of the kinetics and thermodynamics of promoter opening with minimum perturbation.

By comparing the fluorescence of the T7 RNAP−DNA complex of an already melted p-dsDNA with that of dsDNA, we can determine the extent or percentage of open complex at equilibrium. The p-dsDNA is used as a control and as a mimic of a fully melted DNA in an open complex. The p-dsDNA contains an upstream duplex promoter binding sequence (−21 to −5) and a downstream single-stranded initiation and coding sequence (−4 to +19). The t(−4) base in the p-dsDNA is similarly substituted with 2-AP, and although it is unpaired, its fluorescence increases when the p-dsDNA binds to T7 RNAP because the t(−4) base unstacks during open complex formation (*6*, *7*). We have shown that the p-dsDNAs of both ϕ10 and ϕ3.8 promoters bind to T7 RNAP with a very tight affinity in the low nanomolar range (*29*). We assume that the fluorescence intensity of p-dsDNA complexed with T7 RNAP represents the expected fluorescence of a 100% open complex. If all of the dsDNA in complex with T7 RNAP is converted to the open form, then the fluorescence of the ED complex with dsDNA should increase to the same level as the fluorescence of the ED complex with p-dsDNA. We use saturating concentrations of T7 RNAP and DNA to ensure that all of the DNA is bound. As a control, the fluorescence of ED complexes with respective nonfluorescent DNAs was used to correct for any K_d effects and also to eliminate fluorescence contribution of bound T7 RNAP, so that the resulting fluorescence comes only from the 2-AP base in the T7 RNAP−DNA complex (*33*). This method provides a very strict comparison of the 2-AP base environment in dsDNA versus p-dsDNA complexes.

Figure 3A shows the results of such a fluorescence equilibrium experiment. The p-dsDNAs of all three promoters in complex with T7 RNAP showed similar increases in fluorescence. The dsDNAs of all three promoters in complex with T7 RNAP, however, showed a lower increase in fluorescence. Figure 3B shows the ratio of dsDNA complex to p-dsDNA complex fluorescence taken from Figure 3A. Because ϕ10 and ϕ3.8 have different sequences, taking the ratio of dsDNA with respective p-dsDNA rules out any effects of the DNA sequence (such as TAAA versus TATA) on the fluorescence of 2-AP in DNA and thus allows us to compare the percentage or extent of open complex generated by the three promoters. The fluorescence of ϕ10 dsDNA promoter in complex with T7 RNAP is about 30% (28.9 ± 0.6) of the fluorescence of the ϕ10 p-dsDNA complex. Similarly, the fluorescence of the ϕ3.8 dsDNA promoter

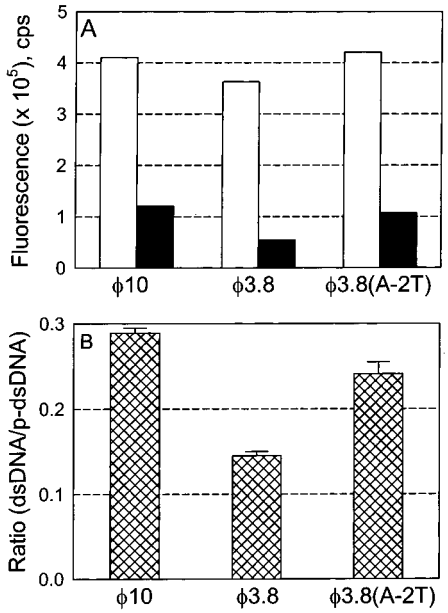


FIGURE 3: Extent of open complex formation. The fluorescence of 40 bp dsDNA or p-dsDNA promoters (1 μM) modified with 2-AP at t(−4) was measured at 25 °C in the presence and absence of T7 RNAP (4 μM). The fluorescence was corrected using the formula fluorescence = F_c(f) − F_c(nf), where F_c(f) and F_c(nf) are the fluorescence intensity of the ED complex with fluorescent (2-AP modified) and nonfluorescent (unmodified) DNAs, respectively. Three measurements were taken for each sample, and the standard error in data acquisition was less than 2% in each experiment. (A) Fluorescence intensity of p-dsDNA (□) and dsDNA (■) in complex with T7 RNAP from a representative experiment. (B) The experiment in (A) was performed twice, and the mean ratio (± standard deviation) of the fluorescence intensity of ED complexes of dsDNA versus p-dsDNA for each promoter DNA is shown. The ratio is 0.289 ± 0.006 for ϕ10, 0.145 ± 0.005 for ϕ3.8, and 0.241 ± 0.014 for ϕ3.8(A−2T) promoters.

complex is 15% (14.5 ± 0.5) of the fluorescence of its p-dsDNA complex, and the fluorescence of the ϕ3.8(A−2T) complex is 25% (24.1 ± 1.4) of the fluorescence of its p-dsDNA complex. These results indicate that the extent of promoter opening in the nonconsensus ϕ3.8 promoter is 50% lower than in the consensus ϕ10 promoter. Interestingly, changing the nonconsensus melting region TAAA to consensus TATA in ϕ3.8(A−2T) results in an increase in the extent of promoter opening to 83% of the consensus TATA sequence. Note that the 2-AP fluorescence provides information about the degree of base unstacking. These results indicate that either the nature of the initiation bubble is different or the equilibrium constant of closed to open complex conversion is different in the three promoters. Assuming that the nature of the initiation bubble is similar with ϕ10 and ϕ3.8 promoters, using the fluorescence values we can estimate the overall equilibrium constant between closed and open complexes. The equilibrium constant, $K_{eq,open}$, for the step $ED_c \leftrightarrow ED_o$ for ϕ10, ϕ3.8, and ϕ3.8-(A−2T) is estimated as 0.3, 0.15, and 0.25, respectively.

The Kinetics of Open Complex Formation Is Affected by the Promoter Sequence. The 2-AP-modified fluorescent promoter DNAs can be used to monitor the real time kinetics of DNA binding and melting. The kinetics of open complex formation under pseudo-first-order conditions was measured by the stopped-flow method. A constant amount of 40 bp DNA with 2-AP at t(−4) was mixed with increasing concentrations of T7 RNAP. The observed rate of promoter opening was obtained from the time-dependent increase in DNA fluorescence, as reported previously (*31*, *33*). The DNA binding kinetics were similar for ϕ10, ϕ3.8, and ϕ3.8(A−2T) promoters. The fluorescence increase with time was monophasic, and the observed rate of this phase showed a hyperbolic dependency on the T7 RNAP concentration for all three promoters (Figure 4).

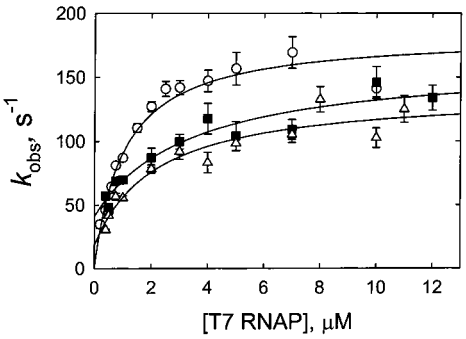


FIGURE 4: Stopped-flow kinetics of open complex formation. The time-dependent increase in fluorescence of t(−4) 2-AP DNA promoter (0.15 μM final) was measured at 25 °C in a stopped-flow instrument at various T7 RNAP concentrations. The observed rate constants, k_{obs}, are plotted against T7 RNAP concentration for ϕ10 (○), ϕ3.8 (■), and ϕ3.8(A−2T) (△). The rate dependency fit to eq 3 with the following $K_{1/2}$, k_{max}, and k_{off} (± standard error) values for ϕ10 (1.0 ± 0.4 μM, 182 ± 14 s^{-1}, and close to zero), ϕ3.8 (3.3 ± 1.9 μM, 120 ± 14 s^{-1}, and 41 ± 11 s^{-1}), and ϕ3.8-(A−2T) (2.0 ± 1.3 μM, 120 ± 12 s^{-1}, and 18 ± 16 s^{-1}).

The minimal mechanism for promoter binding leading to open complex formation can be described by a two-step mechanism, shown in reaction 1:

reaction 1

$$E + D \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ED_c \underset{k_4}{\overset{k_3}{\rightleftharpoons}} ED_o$$

where ED_c is the closed complex and ED_o is the open complex; the forward rate constant, k_3, corresponds to rate of opening (k_{open}), and the reverse rate constant for the same step, k_4, corresponds to rate of closing (k_{close}). This two-step mechanism predicts biphasic kinetics (*39*) where the first phase is the sum of all four intrinsic first-order rate constants ($=k_1[E] + k_2 + k_3 + k_4$) and exhibits a linear dependence on [E], whereas the second phase is expected to show a hyperbolic dependency on [E]. The rates of two phases must differ at least 8−10-fold to be observed as separate phases. Under the conditions of the stopped-flow experiments reported here, the kinetics were monophasic, whose rates increased in a hyperbolic manner with increase in T7 RNAP concentration. The hyperbolic dependency is consistent with the two-step minimal model of open complex formation shown in reaction 1. If the first step, E + D to ED_c conversion, is a rapid equilibrium step, then monophasic kinetics are expected. It is also possible that the first phase is too rapid relative to the time resolution of the stopped-flow instrument and the fluorescence signal for the faster phase is lost in the dead time of the measurement. Alternatively, the first step may not be a rapid equilibrium step, but the rates of the two phases are not very different and hence the two phases are not resolvable.

We attempted to fit the rate dependency data in Figure 4 to the explicit equation for reaction 1 (*39*), but unique values

for intrinsic rate constant k_{open} and k_{close} could not be obtained. We therefore fit the rate dependency to the hyperbolic equation (eq 3), which provided macroscopic parameters such as the apparent k_{off} (y-intercept), $K_{1/2}$, and k_{max} of T7 RNAP−DNA open complex formation. The k_{off} of ϕ10 could not be accurately determined, as the fitting predicted a value close to zero, but ϕ3.8 and ϕ3.8(A−2T) rate dependency indicated that the apparent k_{off} of the nonconsensus promoters is faster and close to 40 and 18 s^{-1}, respectively. The apparent $K_{1/2}$ for DNA binding was close to 1.0 μM for ϕ10 but higher for nonconsensus promoters, equal to 3.3 μM for ϕ3.8 to 2.0 μM for ϕ3.8(A−2T). The meaning of $K_{1/2}$ is not straightforward. If ED_c formation is a rapid equilibrium step, then $K_{1/2}$ corresponds to the K_d of ED_c (k_2/k_1). If ED_c formation step is not a rapid equilibrium step, then $K_{1/2}$ is a complicated function of all the intrinsic rate constants in reaction 1 (*39*). The k_{max} values ranged from 180 s^{-1} for ϕ10 to around 120 s^{-1} for ϕ3.8 and ϕ3.8(A−2T) promoters. The k_{max} is approximately equal to $k_3 + k_4$ or $k_{open} + k_{close}$ (*39*). The apparent k_{on} for ED_o formation can be calculated from the ratio $k_{max}/K_{1/2}$. The consensus promoter has a 3−5-fold higher k_{on} equal to 180 μM^{-1} s^{-1} relative to ϕ3.8 (36 μM^{-1} s^{-1}) and ϕ3.8(A−2T) (60 μM^{-1} s^{-1}). Thus, the nonconsensus promoters form the ED_o complex at an overall slower bimolecular rate constant.

To determine the values of k_{open} and k_{close}, we combine the results of two experiments. The ratio k_{open}/k_{close} or $K_{eq,open}$ was obtained from direct equilibrium fluorescence measurements described in Figure 3, and $k_{max} = k_{open} + k_{close}$ was obtained from the hyperbolic fit of the stopped-flow DNA binding data shown in Figure 4. The k_{open} and k_{close} values for the step $ED_c \Leftrightarrow ED_o$ were thus calculated to be 41 and 141 s^{-1} for ϕ10, 16 and 104 s^{-1} for ϕ3.8, and 23 and 97 s^{-1} for ϕ3.8(A−2T) promoters. These values indicate that although the observed rate of open complex formation ($=k_{max} + k_{off}$) is relatively unaffected by variations in the promoter sequence, the individual rates of opening and closing are affected. Hence the equilibrium constant of the closed to open complex formation step is different and dictated by the promoter DNA sequence.

The Rate of the First Phosphodiester Bond Formation Is Reduced by the Nonconsensus Promoter Sequence. After the promoter DNA is melted and the template is placed at the active site, the initiating NTPs must bind in a template sequence directed manner followed by the formation of the first phosphodiester bond. All three promoters, ϕ10, ϕ3.8, and ϕ3.8(A−2T), initiate with the (+1)GGG sequence; thus the kinetics of initial RNA synthesis can be measured with GTP alone. We have shown for the ϕ10 promoter that the synthesis of pppGpG, the first RNA product, is a rate-limiting step during initiation (*2*, *3*). To determine the effect of promoter sequence on the steps of GTP binding and pppGpG synthesis, the pre-steady-state kinetics of G-ladder RNA synthesis was measured. Previous studies have shown that G-ladder synthesis under pre-steady-state conditions occurs with burst kinetics (*2*, *3*). The fast phase or the burst phase depends on [GTP], and the dependency provides the cumulative K_d of +1 and +2 GTPs and the intrinsic rate of pppGpG synthesis (k_{pol}). High concentrations of T7 RNAP and promoter DNAs enable accurate measurement of the single turnover kinetics. Thus, the K_d of the T7 RNAP−DNA complex or the rate of promoter melting does not affect the observed rate of pppGpG synthesis, although the extent of promoter opening can influence both of these parameters.

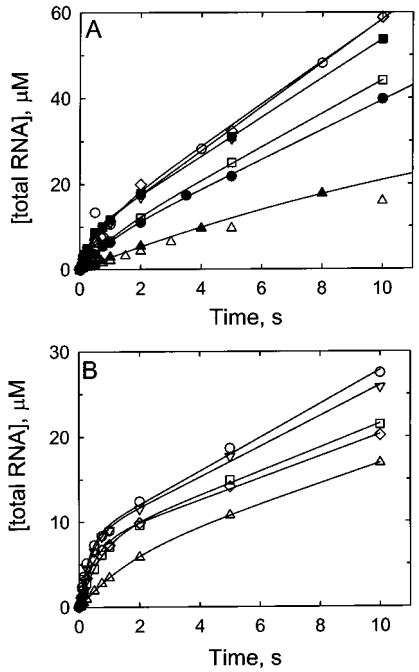


FIGURE 5: Pre-steady-state kinetics of G-ladder RNA synthesis at various GTP concentrations. The time course of G-ladder RNA synthesis by the T7 RNAP−DNA complex (15 μM T7 RNAP + 10 μM DNA, final) and at increasing concentrations of GTP was measured at 25 °C in a rapid chemical quench-flow instrument. (A) shows the pre-steady-state kinetics of total G-ladder RNA synthesis with ϕ3.8 promoter at GTP concentrations of 100 μM (△), 200 μM (▲), 250 μM (●), 300 μM (□), 500 μM (◇), 1000 μM (▽), 1500 μM (○), and 2000 μM (■). (B) shows a similar experiment with ϕ3.8(A−2T) promoter at GTP concentrations of 100 μM (△), 300 μM (□), 500 μM (◇), 1000 μM (▽), and 1500 μM (○).

Figure 5 shows the pre-steady-state kinetics of G-ladder RNA synthesis with ϕ3.8 and ϕ3.8(A−2T) promoters at various [GTP]. The pre-steady-state kinetics of G-ladder RNA synthesis under identical conditions with the ϕ10 promoter has been reported previously (*2*, *3*). All of the promoters show burst kinetics indicating that pppGpG synthesis at the active site of T7 RNAP is fast and a step after pppGpG or G-ladder synthesis, either the product dissociation step or T7 RNAP recycling, is slow. The biphasic kinetics of RNA synthesis with ϕ3.8 is less pronounced relative to ϕ10 or ϕ3.8(A−2T). This is because the steady-state rate of G-ladder synthesis is actually greater, indicating that early RNA products such as pppGpG dissociate faster from the T7 RNAP−ϕ3.8 complex. At 500−600 μM GTP, the observed rate of RNA synthesis with ϕ3.8 is about one-fourth that of ϕ10, and that of ϕ3.8(A−2T) is one-third that of ϕ10. The slower rates of RNA synthesis with the ϕ3.8 promoters relative to ϕ10 could be due to a decrease in the intrinsic rate of initiation, k_{pol}, and/or an increase in the K_d of +1 and +2 GTPs. The pre-steady-state burst rates at various GTP concentrations were analyzed to determine the k_{pol} and K_d values of GTP for each promoter DNA.

The Binding Affinity of Initiating GTPs and Rate of pppGpG Synthesis Are Decreased by the Nonconsensus Promoter Sequence. Figure 6 shows a plot of the pre-steady-state burst rate as a function of [GTP] for all three promoters,

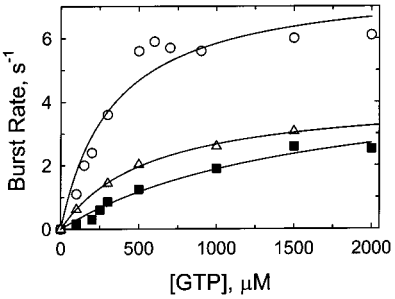


FIGURE 6: Apparent dissociation constant of +1/+2 GTPs for promoters ϕ10, ϕ3.8, and ϕ3.8(A−2T). The biphasic pre-steady-state kinetics of RNA synthesis in Figure 5 were fit to the burst equation (eq 5), and the burst rates are plotted against [GTP]. The [GTP] dependency was fit to the hyperbolic equation (eq 4). The computer fit provided k_{pol} ($=V_{max}$ in eq 4), the maximum rate of pppGpG synthesis (± standard error), equal to 7.8 ± 0.7 s^{-1}, 4.8 ± 0.8 s^{-1}, and 4.2 ± 0.1 s^{-1} and apparent dissociation constants for GTP of 330 ± 86 μM (*3*), 1560 ± 456 μM, and 558 ± 44 μM for promoters ϕ10 (○), ϕ3.8 (■), and ϕ3.8(A−2T) (△), respectively.

ϕ10, ϕ3.8, and ϕ3.8(A−2T). The [GTP] dependencies were fit to the hyperbolic equation to derive the kinetic parameters including the overall K_d of +1 and +2 GTPs and the initiation rate k_{pol}, listed in the table of Figure 8. The pppGpG synthesis occurs with a 2-fold slower rate with the ϕ3.8 promoter, which binds GTP with a 5-fold weaker affinity (1560 ± 456 μM) relative to ϕ10 (330 ± 86 μM) (*3*). The ϕ3.8(A−2T) promoter also shows a close to 2-fold slower rate of pppGpG synthesis, but the affinity for GTP (558 ± 44 μM) is 3-fold tighter than ϕ3.8 and closer in value to ϕ10. The specificity constant, k_{pol}/K_d, is a parameter that provides the overall efficiency of RNA synthesis during initiation, and this parameter is 8 times lower for ϕ3.8 and only 3 times lower for ϕ3.8(A−2T) relative to ϕ10. An interesting result from these studies is that promoters with a consensus −4 to −1 TATA sequence have a higher affinity for initiating GTP. Similarly, promoters with a consensus specificity region sequence show a faster rate of pppGpG synthesis. These effects are most likely due to allosteric mechanisms and not the result of the different binding affinities of the various promoters. Thus, promoter specificity and melting regions that are away from the site of RNA synthesis can affect the efficiency of RNA synthesis by modulating different steps of initiation.

The Binding Affinity of +1 GTP Is Decreased by the Nonconsensus Promoter Sequence. Jia and Patel (*3*) have shown that the kinetics of GTP binding can be measured by following the fluorescence of 2-AP in promoter DNA. The 2-AP probe can be incorporated in place of adenines in the −4 to +4 region, and modification at any of these positions provides a sensitive fluorescence signal for measuring the kinetics of GTP binding. In the experiments described here, a complex of T7 RNAP and nt(+4) 2-AP-modified promoter (ϕ3.8 and ϕ10) was mixed with GTP in a stopped-flow instrument. The fluorescence of 2-AP at nt(+4) increases in a time-dependent manner upon GTP binding as shown in Figure 7A. The observed exponential rate increases with increasing [GTP] as shown in Figure 7B, and the K_d of +1/+2 GTP can be derived from the hyperbolic dependencies. Consistent with the radiometric RNA synthesis assay (Figure 6), the overall K_d of +1 and +2 GTPs is 3-fold higher with the ϕ3.8 promoter and the rate of a conformational change induced by GTP binding is 2.5-fold slower relative to ϕ10.

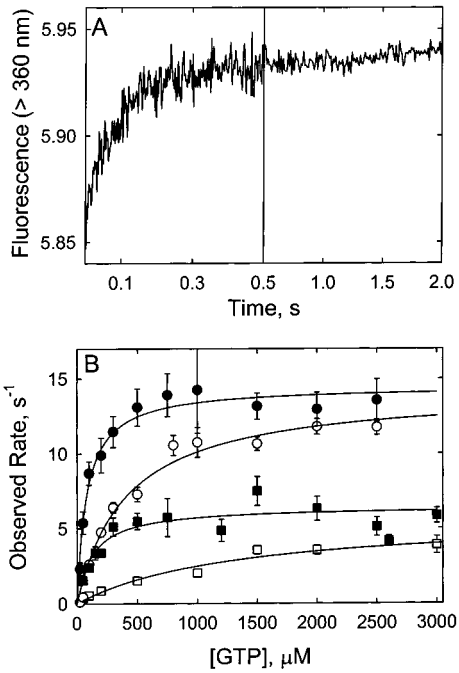
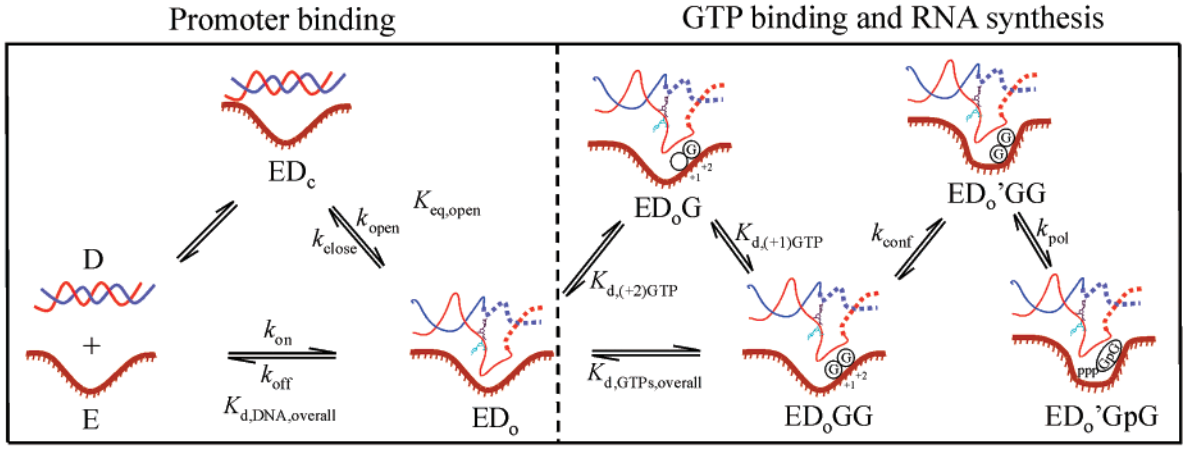


FIGURE 7: Stopped-flow kinetics of initiating GTP binding. (A) The fluorescence of nt(+4) 2-AP DNA (0.15 μM final) in complex with T7 RNAP (0.45 μM final) increases as a function of time upon addition of GTP. The GTP binding kinetics at 25 °C were measured at various [GTP], and a representative trace (in two time windows) at 2500 μM GTP is shown for ϕ10 promoter. (B) The observed exponential rate of fluorescence increase was plotted against [GTP], both in the absence of GMP and in the presence of GMP. The data were fit to the hyperbolic equation (eq 4), which provided an apparent K_d (± standard error) for ϕ10 and ϕ3.8 promoters of 400 ± 70 μM (○) and 1300 ± 275 μM (□) in the absence of GMP and 80 ± 10 μM (●) and 130 ± 50 μM (■) in the presence of 600 and 1000 μM GMP, respectively. The maximum rate of conformational change induced by GTP binding is 14.2 ± 0.8 s^{-1} for ϕ10 promoter and 5.7 ± 0.4 s^{-1} for ϕ3.8 promoter, in the absence of GMP. The corresponding rates in the presence of GMP are 14.5 ± 0.4 s^{-1} for ϕ10 promoter and 6.4 ± 0.5 s^{-1} for ϕ3.8 promoter.

The same experiment was then carried out in the presence of a constant amount of GMP to determine the relative K_d values of +1 and +2 GTPs. It has been shown that T7 RNAP can initiate very efficiently with GMP (*40*), implying that the +1 position can bind GMP. However, the binding of GMP alone does not show any fluorescence change, but the addition of GTP provides the observed fluorescence change. Thus, in the presence of a saturating concentration of GMP, one can estimate the K_d of +2 GTP. In the presence of GMP, the apparent K_d of GTP is tight, and ϕ3.8 shows only a 1.5-fold higher K_d for +2 GTP relative to ϕ10. From the relationship $K_{d,GTPs,overall} = (K_{d,(+1)GTP}K_{d,(+2)GTP})^{1/2}$, we calculate the K_d of +1 GTP close to 13 mM for ϕ3.8 and 2 mM for ϕ10 promoter. These results indicate that the affinity of +1 GTP is modulated by the promoter sequence, but the affinity of +2 GTP appears to be less sensitive to the promoter sequence. Thus, GTP concentration can differentially regulate the efficiency of transcription initiation at various promoters.

DISCUSSION

T7 RNAP does not require any accessory protein factors for initiation, and the efficiency of transcription is regulated



Constants	ϕ10	ϕ3.8	ϕ3.8(A-2T)
[1]$K_{d,DNA,overall}$ (μM)	0.47 ± 0.02	0.95 ± 0.10	1.13 ± 0.08
k_{on} (μM^{-1} s^{-1})	182	36	60
k_{off} (s^{-1})	~ 0	41 ± 11	18 ± 16
[2]$K_{eq,open}=k_{open}/k_{close}$	0.29 ± 0.01	0.15 ± 0.01	0.24 ± 0.01
$k_{max}=k_{open}+k_{close}$($s^{-1}$)	182 ± 14	120 ± 14	120 ± 12
k_{open} (s^{-1})	41	16	23
k_{close} (s^{-1})	141	104	97

Constants	ϕ10	ϕ3.8	ϕ3.8(A-2T)
[3]$K_{d,GTPs,overall}$ (μM)	330 ± 86	1560 ± 456	558 ± 44
[4]$K_{d,GTPs,overall}$	400 ± 70	1300 ± 275	nd
[4]$K_{d,(+2)GTP}$ (μM)	80 ± 10	130 ± 50	nd
[4]$K_{d,(+1)GTP}$ (mM)	2	13	nd
[4]k_{conf} (s^{-1})	14.2 ± 0.8	5.7 ± 0.4	nd
[3]k_{pol} (s^{-1})	7.8 ± 0.7	4.8 ± 0.8	4.2 ± 0.1
k_{pol}/K_d ($s^{-1}\mu M^{-1}$)	0.024	0.003	0.008

FIGURE 8: Transcription initiation pathway of consensus and nonconsensus T7 promoters. The minimal pathway of transcription initiation defined by studies described here and in previous reports (*2*, *3*, *29*, *31*) is illustrated. T7 RNAP and 40 bp dsDNA promoter are represented by E and D, respectively. The template and nontemplate strands of DNA are colored red and blue, respectively. Part of the DNA shown in solid lines in the ED$_o$ complexes was constructed using the DNA coordinates from the crystal structure of the initiation complex (PDB id 1QLN). The dotted line is the hypothetical extension of the two strands that illustrate the transcription bubble during initiation. The table compares the kinetic and equilibrium parameters of each step during initiation for the three promoters obtained from the following measurements: (1) fluorometric titrations (Figure 2), (2) equilibrium fluorescence measurements (Figure 3), (3) pre-steady-state kinetics of RNA synthesis (Figure 6), and (4) stopped-flow GTP binding (Figure 7). Uncertainties of only experimentally determined parameters are reported. The parameters such as apparent k_{on}, k_{open}, k_{close}, $K_{d,(+1)GTP}$, and k_{pol}/K_d calculated from experimental values are reported without uncertainties.

by the T7 promoter sequence (and defined as promoter strength), but the kinetic and thermodynamic basis has not been established. In this paper, we have investigated the kinetics and thermodynamics of each step of transcription initiation by T7 RNAP using promoter sequences to understand the effect of the divergent sequence on promoter strength relative to the consensus sequence. Rationally speaking, any number of sequences from class II and class III natural T7 promoter selection could be used for promoter strength comparison. Since all five class III strong promoters have absolutely conserved sequence from −17 to +6, only one promoter from this class is sufficient as an upper limit of promoter strength, and for our studies we chose the *ϕ*10 promoter. Of the 10 class II weak promoters, only five promoters (*ϕ*1.1A, *ϕ*1.3, *ϕ*3.8, *ϕ*4c, and *ϕ*4.7) have divergent bases in the promoter binding domain (−17 to −5), of which only one promoter, i.e., *ϕ*3.8, has the most nonconsensus bases (total three bases) in the promoter specificity region (−13 to −6). The *ϕ*3.8 promoter also has an altered base (at −2) in the melting region (−4 to −1) among six class II promoters that have the most nonconsensus bases (one base) in this region. Thus, with most divergent bases in both promoter specificity and melting regions, the *ϕ*3.8 promoter represents the lower limit of promoter strength for comparison. An artificial class II promoter *ϕ*3.8(A−2T) divides these upper and lower limits. Altogether, these three promoters provide sequences for "distinct" comparison of the effect of the specificity region only (*ϕ*10 vs *ϕ*3.8(A−2T)) or the melting region only (*ϕ*3.8 vs *ϕ*3.8(A−2T)) or both (*ϕ*10 vs *ϕ*3.8). It is interesting to note that the 10 divergent promoters from class II constitute an average of four nonconsensus bases per promoter, of which three lie in the initiation domain and one in the binding domain. Therefore, it is not surprising that regulation at the level of transcription initiation occurs mainly from the initiation domain and to a lesser extent from the promoter binding domain. Transcription initiation is a multistep process, as shown in Figure 8, and in principle each step is a target of transcription regulation. Transcription can be regulated at the level of promoter recognition, open complex formation, by the affinity of initiating nucleotides, and the rate constants of initiation, promoter clearance, and processivity during abortive and elongation phases of RNA synthesis. The rate-limiting steps and those with unfavorable equilibrium constants have a greater influence on the overall efficiency of transcription. Our goal in these studies was to identify the steps of initiation that are most affected by the promoter sequence. The results of our studies with three promoters are summarized in Figure 8.

The first step during transcription initiation is the specific recognition of the promoter by the T7 RNAP, resulting in the formation of a specific closed complex in which an initiation bubble is not yet formed. The closed complex

isomerizes to an open complex in which the TATA sequence from −4 to −1 region is melted (*31*). With the 40 bp promoters, the resulting open complex would contain a bubble approximately in the middle of the DNA. Using 2-AP DNA and stopped-flow transient state kinetics, we infer that the ϕ3.8 promoter forms a much weaker closed complex than the ϕ10 promoter. The equilibrium fluorescence measurements indicate a 2-fold reduction in binding affinity of ϕ3.8 promoters. Thus, bases at positions −11, −12, and −13 are primary determinants of the binding affinity during closed complex formation. The crystal structure shows specific interaction of guanine at nt(−11) of the ϕ10 promoter sequence with the side chain of N748 (*6*, *7*). Similarly, nt(−13) and nt(−14) bases interact respectively with the side chain amine and backbone amide of K98 (from the AT-rich recognition loop). The relative K_d values of ϕ3.8 and ϕ10 promoters indicate that these interactions correspond to a $\Delta\Delta G$ value of 0.4−0.6 kcal/mol.

The overall on-rate, k_{on}, for the three promoters indicates that nonconsensus promoters bind T7 RNAP at a rate 3−5-fold slower than the consensus promoter. Similarly, the overall k_{off} values for the three promoters suggest that nonconsensus promoters form a kinetically unstable complex with T7 RNAP. These results indicate that T7 RNAP binds more efficiently and also stays more stably bound to a consensus promoter. Several studies indicate that the formation of open complex in T7 RNAP (in the absence of MgGTP) is an unfavorable equilibrium reaction. Thus, the preinitiation ED complexes consist of a mixture of ED_c and ED_o. The ratio of ED_c to ED_o is estimated to be between 7:1 by steady-state RNA synthesis assay (*34*) and 3:1 by the relative fluorescence increase of dsDNA versus p-dsDNA in ϕ10 promoter DNA complexes with T7 RNAP (*31*). The fluorescence method therefore provides a $K_{eq,open}$ for the ϕ10 promoter around 0.3, indicating that at equilibrium 30% of the total ED complexes is in the open form. Similar measurements with the ϕ3.8 promoter provided a $K_{eq,open}$ of 0.15, indicating that only 15% of the total ED complexes exists as open complexes. Interestingly, the ϕ3.8(A−2T) promoter with a consensus melting region showed a higher amount of open complex (25%), almost close to the level observed in the ϕ10 promoter. Thus, kinetic and thermodynamic studies of promoter binding and opening indicate that base changes in the promoter specificity region affect the overall stability of the ED complexes whereas base changes in the melting region affect mainly the stability of the ED_o complex. The observed rates of open complex formation ($=k_{max} + k_{off}$) are relatively insensitive to base changes in the specificity and/or melting region, but the individual rates, k_{open} and k_{close}, are affected, and both nonconsensus promoters show a slower k_{open}. The values of k_{open} and k_{close} are determined by considering a minimum two-step mechanism; however, a complete mechanism of open complex formation may involve more than two steps. Therefore, intrinsic rates of promoter opening and closing need to be determined either by direct measurements or by computer simulation of the kinetic data.

How can a single base change in the TATA region affect the stability of ED_o and the equilibrium between closed and open complexes? One parameter that can have a global effect on the stability of the melted DNA is the bendability or deformability of the sequence in the melting region. It has been proposed that DNA bending is the first step that occurs during promoter opening, and bending the DNA results in base unstacking and nucleation of DNA melting at the RNAP active site (*41*). T7 promoters with TATA and TGCA sequences in the melting region contain an intrinsic bend (*42*). When these promoters bind T7 RNAP, the DNA is further bent 40−60°, and the bending is postulated to be centered at the −2 to +1 position (*42*). When we examine the trinucleotide parameters that predict bendability of the −4 to −1 region of the dsDNA, they indicate that the bendability of the −4 to −1 region with sequence TATA and TGCA is much higher than that with TAAA (*43*). Thus, the TATA box, a ubiquitous sequence found in promoters, may be intrinsically easier to bend as well as easier to maintain in a single-stranded form relative to other sequences such as the TAAA sequence. Thus, a single base change can alter the local bendability of the dsDNA, making nucleation of DNA melting an unfavorable process and influencing on multiple steps of initiation.

The promoter sequence also influences the K_d of +1 and +2 GTPs as well as the observed rate of the first phosphodiester bond, the pppGpG synthesis. This observation is similar to that of *E. coli* rrn promoters that appear to be regulated by the initiating nucleotide concentration (*44*). We show here that T7 promoters are differentially regulated by initiating NTP concentration. The pre-steady-state kinetics of transcription initiation showed that nonconsensus promoters have an overall weaker affinity for initiating GTPs and a slower rate of pppGpG synthesis. More specifically, it is the K_d of +1 GTP that limits the rate of pppGpG synthesis. A promoter with a consensus melting sequence (TATA) binds GTP with an overall tighter affinity relative to a promoter with a nonconsensus melting region (TAAA). The initiation rate, that is, the rate of pppGpG formation, on the other hand is dependent on the sequence of the specificity region. A promoter with a nonconsensus specificity region synthesizes pppGpG with a 2−3-fold slower rate relative to a consensus promoter. Since the specificity region is away from the site of RNA synthesis, we propose that conformational changes resulting from the interactions between the T7 RNAP and the specificity region are somehow relayed to the site of catalysis to modulate the rate of initiation.

CONCLUSIONS

The comparative studies of ϕ10 and ϕ3.8 promoters shed light on the transcription initiation steps that are sensitive to the promoter sequence. These studies have identified some of the "intrinsic" regulatory mechanism built into the promoter sequence that the T7 RNAP enzyme uses during initiation to control the level of RNA synthesis without the aid of other protein factors. The results indicate that multiple steps in the pathway of transcription initiation are affected by the promoter sequence. Some of these steps include the stability of the closed and open complexes, the rate of the first phosphodiester bond formation, and the affinity for the +1 initiating nucleotide. Transcription control at the promoter melting step occurs by regulation of the equilibrium constant between closed and open complexes and to a lesser extent by the rate of open complex formation. Our studies also indicate that although base variations in the promoter sequence affect the individual steps to a small degree, the

cumulative effect can significantly affect the overall yield of the final transcript and dictate the overall promoter strength.

REFERENCES

1. deHaseth, P. L., Zupancic, M. L., and Record, M. T., Jr. (1998) *J. Bacteriol. 180*, 3019−3025.
2. Jia, Y., and Patel, S. S. (1997) *Biochemistry 36*, 4223−4232.
3. Jia, Y., and Patel, S. S. (1997) *J. Biol. Chem. 272*, 30147−30153.
4. Cheetham, G. M., Jeruzalmi, D., and Steitz, T. A. (1998) *Cold Spring Harbor Symp. Quant. Biol. 63*, 263−267.
5. Cheetham, G. M., and Steitz, T. A. (2000) *Curr. Opin. Struct. Biol. 10*, 117−123.
6. Cheetham, G. M. T., and Steitz, T. A. (1999) *Science 286*, 2305−2309.
7. Cheetham, G. M. T., Jeruzalmi, D., and Steitz, T. A. (1999) *Nature 399*, 80−83.
8. Rong, M., He, B., McAllister, W. T., and Durbin, R. K. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 515−519.
9. Gross, L., Chen, W. J., and McAllister, W. T. (1992) *J. Mol. Biol. 228*, 488−505.
10. Joho, K. E., Gross, L. B., McGraw, N. J., Raskin, C., and McAllister, W. T. (1990) *J. Mol. Biol. 215*, 31−39.
11. Raskin, C. A., Diaz, G., Joho, K., and McAllister, W. T. (1992) *J. Mol. Biol. 228*, 506−515.
12. Lee, S. S., and Kang, C. (1993) *J. Biol. Chem. 268*, 19299−19304.
13. Klement, J. F., Moorefield, M. B., Jorgensen, E., Brown, J. E., Risman, S., and McAllister, W. T. (1990) *J. Mol. Biol. 215*, 21−29.
14. Dunn, J. J., and Studier, F. W. (1983) *J. Mol. Biol. 166*, 477−535.
15. McAllister, W. T., and Carter, A. D. (1980) *Nucleic Acids Res. 8*, 4821−4837.
16. Ikeda, R. A., Lin, A. C., and Clarke, J. (1992) *J. Biol. Chem. 267*, 2640−2649.
17. Ikeda, R. A. (1992) *J. Biol. Chem. 267*, 11322−11328.
18. Ikeda, R. A., Ligman, C. M., and Warshamana, S. (1992) *Nucleic Acids Res. 20*, 2517−2524.
19. Kumar, A., and Patel, S. S. (1997) *Biochemistry 36*, 13954−13962.
20. Zhang, X., and Studier, F. W. (1997) *J. Mol. Biol. 269*, 10−27.
21. Ikeda, R. A., and Bailey, P. A. (1992) *J. Biol. Chem. 267*, 20153−20158.
22. Villemain, J., and Sousa, R. (1998) *J. Mol. Biol. 281*, 793−802.
23. Huang, J., Villemain, J., Padilla, R., and Sousa, R. (1999) *J. Mol. Biol. 293*, 457−475.
24. Chapman, K. A., and Burgess, R. R. (1987) *Nucleic Acids Res. 15*, 5413−5432.
25. Diaz, G. A., Raskin, C. A., and McAllister, W. T. (1993) *J. Mol. Biol. 229*, 805−811.
26. McClure, W. R. (1985) *Annu. Rev. Biochem. 54*, 171−204.
27. Gralla, J. D. (1996) *Methods Enzymol. 273*, 99−110.
28. Cermakian, N., Ikeda, T. M., Cedergren, R., and Gray, M. W. (1996) *Nucleic Acids Res. 24*, 648−654.
29. Jia, Y. P., Kumar, A., and Patel, S. S. (1996) *J. Biol. Chem. 271*, 30451−30458.
30. Grodberg, J., and Dunn, J. J. (1988) *J. Bacteriol. 170*, 1245−1253.
31. Bandwar, R. P., and Patel, S. S. (2001) *J. Biol. Chem. 276*, 14075−14082.
32. Ujvari, A., and Martin, C. T. (1997) *J. Mol. Biol. 273*, 775−781.
33. Ujvari, A., and Martin, C. T. (1996) *Biochemistry 35*, 14574−14582.
34. Villemain, J., Guajardo, R., and Sousa, R. (1997) *J. Mol. Biol. 273*, 958−977.
35. Sasse-Dwight, S., and Gralla, J. D. (1989) *J. Biol. Chem. 264*, 8074−8081.
36. Brieba, L. G., and Sousa, R. (2001) *Biochemistry 40*, 3882−3890.
37. Rachofsky, E. L., Osman, R., and Ross, J. B. A. (2001) *Biochemistry 40*, 946−956.
38. Ward, D. C., Reich, E., and Stryer, L. (1969) *J. Biol. Chem. 244*, 1228−1237.
39. Johnson, K. A. (1992) Transient-state kinetic analysis of enzyme reaction pathways, *Enzymes (3rd Ed.) 20*, 1−61.
40. Martin, C. T., and Coleman, J. E. (1989) *Biochemistry 28*, 2760−2762.
41. Travers, A. A. (1990) *Cell 60*, 177−180.
42. Ujvari, A., and Martin, C. T. (2000) *J. Mol. Biol. 295*, 1173−1184.
43. Goodsell, D. S., and Dickerson, R. E. (1994) *Nucleic Acids Res. 22*, 5497−5503.
44. Gaal, T., Bartlett, M. S., Ross, W., Turnbough, C. L., and Gourse, R. L. (1998) *Science 278*, 2092−2097.

BI0158472