Spectroscopic Determination of Open Complex Formation at Promoters for Escherichia coli RNA Polymerase †

ARTICLE *in* BIOCHEMISTRY · AUGUST 1997

Impact Factor: 3.02 · DOI: 10.1021/bi970363k · Source: PubMed

4 AUTHORS, INCLUDING:

Keith Bjornson
10X Genomics
18 PUBLICATIONS **2,352** CITATIONS
SEE PROFILE

Pieter L. DeHaseth
Case Western Reserve University
73 PUBLICATIONS **4,085** CITATIONS
SEE PROFILE

Spectroscopic Determination of Open Complex Formation at Promoters for *Escherichia coli* RNA Polymerase†

Jennifer J. Sullivan,‡ Keith P. Bjornson,§ Lawrence C. Sowers,‖ and Pieter L. deHaseth*,‡

Department of Biochemistry, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106-4935, Department of Biochemistry and Molecular Biophysics, Box 8231, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, Missouri 63110, and Division of Pediatrics, City of Hope National Medical Center, Duarte, California 91010

Received February 18, 1997; Revised Manuscript Received May 6, 1997⊗

ABSTRACT: A considerable amount of effort has been expended studying the kinetics of association of *Escherichia coli* RNA polymerase with promoter DNA in forming the open complex. Strand separation occurs over about 12 base pairs and includes the transcription start site. However, these efforts have been significantly hampered by the lack of a sensitive, real time method by which formation of an open complex could be assayed. Here, we employ short (86 bp) synthetic promoters with 2-aminopurine (2-AP) substitutions in the region that becomes single-stranded to spectroscopically monitor open complex formation. We demonstrate that promoters bearing the substitutions behave in a manner similar to that of those containing only the four common bases with respect to both the region of strand separation and start site selection. Open complex formation was found to yield an increased fluorescence signal with an emission maximum characteristic of 2-aminopurine. This spectroscopic assay for open complex formation was found to be well-suited to the investigation of a strong promoter, allowing open complex formation to be followed over a time scale of seconds with a stopped flow apparatus. The introduction of two additional nonconsensus base pairs in the −35 region resulted in a promoter for which open complex formation was 100-fold slower. The same substrates were also used to monitor the promoter re-annealing that ensues upon initiation of RNA synthesis. Similar rates for this process were observed for the two promoter variants employed in this study.

Control of transcription initiation can be exerted by modulating the kinetics of formation of open RNA polymerase−promoter complexes (*1*). Thus, a considerable amount of effort has been expended in studying the kinetics of association of *Escherichia coli* RNA polymerase with DNA. However, these efforts have been significantly hampered by the lack of a sensitive, real time method by which formation of an open complex could be assayed. Here, we demonstrate that the use of promoters with 2-aminopurine (2-AP)[1] substitutions in the region that becomes single-stranded provides just such a method. The availability of a spectroscopic signal opens the study of the interaction of *E. coli* RNA polymerase with its promoters to the use of conventional stopped flow techniques, readily enabling the determination of fast rates of open complex formation.

2-Aminopurine is an analog that forms a Watson−Crick base pair with T in a normal B-form DNA helix which is only slightly weaker than an A-T base pair (*2*−*5*). In a DNA template, 2-AP pairs predominantly with dTTP during polymerase-directed DNA synthesis and oligodeoxynucleotides containing 2-AP within the recognition site of *Eco*RI restriction endonuclease are cleaved, albeit at a reduced rate (*6*). However, unlike the four common DNA bases, 2-AP is fluorescent at neutral pH (*7*), and it has been demonstrated that its fluorescence is sensitive to the local environment. Upon formation of a well-stacked base pair, the 2-AP fluorescence is attenuated (*8*). 2-AP fluorescence can therefore be used as a nonperturbing probe of local DNA conformational changes (*9*).

Classically, two DNA elements have been identified as being important for the recognition of promoter DNA by *E. coli* RNA polymerase: hexameric conserved regions at −35 and −10 with consensus sequences TTGACA and TATAAT, respectively on the nontemplate strand (*10*). Some promoters have a third element, an A + T rich region between −40 and −60 (*11*), but this is not the case for the P_{RM} promoter (*12*), variants of which are the subject of the current study. In a functional or open complex, a stretch of DNA extending from the middle of the −10 region to just past the start site of initiation is rendered single-stranded by the RNA polymerase (*10*). By substitution of 2-AP into this region, the melting can be monitored spectroscopically. While our studies were in progress, the use of 2-AP-substituted DNA was reported for the study of the interaction of both *E. coli* RNA polymerase (*13*) and T7 RNA polymerase with their promoters (*14*−*16*).

MATERIALS AND METHODS

Materials. RNA polymerase was purified as described (*17*, *18*). Titration of 2-AP-substituted DNA **C** (see Figure

† We gratefully acknowledge support from NIH Grants GM31808 (P.L.dH.), GM41336 (L.C.S.), T32GM08492 (K.P.B.), and GM45948 (to T. Lohman).

* Address all correspondence to this author at Department of Biochemistry, School of Medicine, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106-4935. E-mail: dehaseth@biocserver.cwru.edu. Fax: (216) 368-3419.

‡ Case Western Reserve University.

§ Washington University School of Medicine.

‖ City of Hope Natinoal Medical Center.

⊗ Abstract published in *Advance ACS Abstracts,* June 15, 1997.

[1] Abbreviations used: RNAP, RNA polymerase; bp, base pair; 2-AP, 2-aminopurine.

S0006-2960(97)00363-2 CCC: $14.00

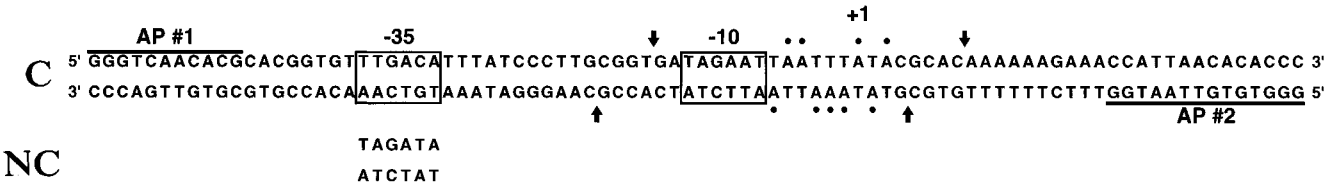

FIGURE 1: Promoter fragments used in this work. The 86 bp fragments were assembled from six oligodeoxynucleotides, the junctions of which are indicated by arrows. Positions of 2-AP substitutions (a total of nine in the double helical DNA) are indicated by black dots. The two variants of the P_{RM} promoter differ in two positions of the −35 region, as shown: **NC** has a nonconsensus −35 region, and **C** a consensus sequence in this region. The sequences of two primers used in the experiments described in Figures 2 and 3 are also indicated.

1) with RNA polymerase (data not shown) showed an equivalence point of approximately two RNA polymerase molecules per promoter DNA, consistent with a 50% activity of the RNA polymerase in our preparation. This is similar to that found by other methods. Polynucleotide kinase, DNA ligase, and Klenow fragment of DNA polymerase I were from Amersham/USB; they were used per the manufacturer's recommendation. Polyacrylamide/urea sequencing gel mix was from Sequagel.

2-AP-Containing Oligodeoxynucleotides. Oligodeoxynucleotides containing 2-AP substitutions were synthesized as recently described (*19*).

Promoter Construction. Promoter assembly followed established procedures (*20*). Starting with 1 nmol of oligodeoxynucleotide, the yield of the gel-purified assembled promoter fragment was 200−400 pmol.

Molecular Biology Techniques. Probing with potassium permanganate was used to determine the region of strand separation in open complexes. The previously employed procedure (*21*) was used with some modifications. After incubation of RNA polymerase (200 nM) and promoter fragment (20 nM) for 15 min in transcription buffer [30 mM HEPES (pH 7.6), 100 mM KCl, 10 mM $MgCl_2$, and 0.1 mM DTT] with 50 μg/mL BSA, $KMnO_4$ was added to 46 mM, and after 2 min, the reaction was quenched with β-mercaptoethanol. The DNA was purified by phenol/chloroform extraction and passage through a Sephadex G-50 spin column. ^{32}P-labeled primer (to 20 nM) was added and hybridized by heating and cooling; after extension with Klenow DNA polymerase, the reaction was analyzed on a sequencing gel.

Determination of the start site by primer extension was as described (*22*). RNA was synthesized in a reaction with a final RNAP and promoter fragment concentration of 100 nM each in transcription buffer containing 50 μg/mL BSA. The unlabeled RNA was purified by electrophoresis on a polyacrylamide/urea gel. Primer (AP 2) extension reaction mixtures each contained 20% of the thus-purified RNA.

Fluorescence Assays. All experiments were carried out on a Perkin-Elmer LS50B spectrofluorometer using a 300 μL micro cell. The sample compartment was thermostated at the desired temperature. Samples were excited at 310 nm (5 nm slit width), and emission was monitored at 370 nm (15 nm slit width). For kinetic experiments, binding reactions were initiated by addition of active RNA polymerase in sufficient excess over the promoter DNA (10 nM) so that pseudo-first-order conditions would apply (see below). In general, 160 μL aliquots of RNA polymerase and DNA at twice the desired final concentrations in transcription buffer were preincubated at the temperature of the experiment and mixed in the cuvette at time 0. As the RNA polymerase was stored in 50% glycerol, adjustments were made to obtain a final glycerol concentration of 10% in all reactions mixtures. Readings were taken at intervals of 15 or 30 s for 5 (**C** template) or 15 (**NC** template) min.

Measurements of the loss of fluorescence signal were carried out on RNA polymerase−DNA complexes that had been preincubated for 10−15 min to ensure formation of open complexes. A mixture of the four NTPs and heparin was then added to final concentrations of 200 mM and 50 μg/mL, respectively. The fluorescence intensity was measured at 1 min intervals for 8−12 min.

Stopped Flow Experiments. The experiments were carried out on an Applied Photophysics SX17MV stopped flow apparatus in the fluorescence mode. The excitation monochromator was set at 315 nm, and a 350 nm "cut-on" filter (transparent to light with a wavelength of greater than 350 nm) was used for the emitted light. The syringes were loaded with samples prepared exactly as indicated above for the steady state fluorescence measurements, except that the volumes were adapted to the needs of each experiment. Each shot involved the mixing of 55 μL volumes of DNA and RNA polymerase solutions in a thermostated reaction cell; during the interval in which the progress of the reaction was followed, 1000 data points were collected. For each determination of k_{obs}, the values from three or more shots were averaged.

Data Analysis. Data analysis was carried out with Sigma plot 4.17 (Jandel Scientific). The experiments were usually conducted with RNA polymerase in an at least 3-fold excess over the promoter fragments so that the kinetics would be pseudo-first-order. The data were then fit to the equation $y = a[1 - \exp(-k_{obs}t)] + c$, where y is the fluorescence signal at 370 nm, t is the time in seconds after mixing promoter DNA and RNA polymerase, k_{obs} is the pseudo-first-order rate constant, a is the observed amplitude, and c is the starting fluorescence signal. Under any set of conditions, k_{obs} was obtained at several RNA polymerase concentrations. The parameters K_B and k_f (see Scheme 1 in Results) as well as the second-order rate constant (as $k_a = K_B k_f$) were obtained from the hyperbolic fit of the data to the equation $k_{obs} = k_a[\text{RNAP}]/(K_B[\text{RNAP}] + 1)$. The double-reciprocal form $k_{obs}^{-1} = k_f^{-1} + k_a^{-1}\,[\text{RNAP}]^{-1}$ was used if it became clear that the data were not of sufficient quality to permit a hyperbolic fit (see Results and the footnotes of Table 1). In the latter case, the second-order rate constant k_a was evaluated more appropriately from the slopes of plots of k_{obs} as a function of [RNAP].

RESULTS

P_{RM} promoter variants similar to those employed in this study (Figure 1) have been previously characterized (*23*). The 86 bp fragments were assembled from six oligodeoxy-

Table 1: Summary of Kinetic Data for **C** and **NC**

	C			**NC**
temp (°C)	k_f (s^{-1})	K_B (M^{-1})	k_a ($M^{-1}\,s^{-1}$)	k_a ($M^{-1}\,s^{-1}$)
37	2.5 ± 0.7[a]	$(0.8 \pm 0.3) \times 10^7$ [a]	$(2.0 \pm 0.5) \times 10^7$ [a]	$(1.6 \pm 0.2) \times 10^5$ [b]
30	–	–	–	$(1.2 \pm 0.3) \times 10^5$ [b]
25	0.4−9[c]	$(0.1-2) \times 10^7$ [c]	$(0.6 \pm 0.1) \times 10^7$ [c]	$(0.6 \pm 0.1) \times 10^5$ [b]
15	–[b]	–[b]	$(0.3 \pm 0.02) \times 10^7$ [b]	–

[a] The data of figure 6 were fit to the equation $k_{obs} = k_a[\text{RNAP}]/(K_B[\text{RNAP}] + 1)$; error estimated from this as well as from the double-reciprocal fit ($k_{obs}^{-1} = k_f^{-1} + k_a^{-1}[\text{RNAP}]^{-1}$). [b] k_a was evaluated from the slope of a plot of k_{obs} versus [RNAP], constrained to go through the origin; the data did not permit deconvolution into the components K_B and k_f. [c] Double-reciprocal fit. Range of values based on the range found for k_f from the fitting procedure.

nucleotides (as delineated by the arrowheads in the figure) following established procedures (*20*). With the goal of developing a spectroscopic assay for open complex formation, we have introduced 2-AP substitutions for A residues in the promoter region that becomes single-stranded upon open complex formation. The middle oligodeoxynucleotides in Figure 1 each have several substitutions (as indicated by the black dots) for a total of nine 2-AP in the double-stranded DNA, in order to maximize the signal. No substitutions were made in the −10 region to minimize the possibility of the modified nucleotide interfering with direct polymerase−DNA interactions. The ligation of the "outside" oligonucleotide pairs to that containing the 2-AP yielded the full length DNA fragment with an overall efficiency of 20−40%. In view of the relatively small amounts of DNA used in each experiment (about 1 pmol of DNA), even a small scale (40 nmol) synthesis of 2-AP-substituted DNA would suffice for the purpose of assembling enough template for a large number of experiments. The experiments reported here were carried out on two promoter variants: variant **C** has the consensus and variant **NC** the wild type −35 region of the P_{RM} promoter (see Figure 1). Our previous studies (*23*) showed that *in vitro* open complex formation at **C** proceeded with a rate constant that was about 3 orders of magnitude greater than at **NC** and that *in vivo* **C**-driven expression of a reporter gene resulted in 5-fold more product.

The use of chemically synthesized DNA fragments (necessary for the base analog substitutions) imposes length constraints on the promoters employed in these studies. The main disadvantage of a promoter sequence embedded in a short DNA fragment is the potential for its occlusion by RNA polymerase molecules bound at either end. RNA polymerase has a propensity for end binding (*24*), possibly because the fraying of the strands creates stretches of single-stranded DNA for which the enzyme has high affinity. We attempted to circumvent this problem through the use of "G-C clamps" at both ends of the fragment (see Figure 1). Because of this potential problem and the unknown effect of the 2-AP substitutions on open complex formation, it was important to determine whether RNA polymerase interacted properly with the promoters on the fragments, regardless of the 2-AP substitutions.

In Figure 2, we present evidence that the region of RNA polymerase-induced strand separation is as expected for the P_{RM} promoter. Oxidation by $KMnO_4$ preferentially targets T residues in single-stranded DNA (*25*), as would result from RNA polymerase-induced strand separation in the open complex. Modified T residues were detected by primer extension (*25*); DNA synthesis comes to a halt when the template presents a modified T at the position that specifies the next nucleotide to be incorporated. The modified DNA

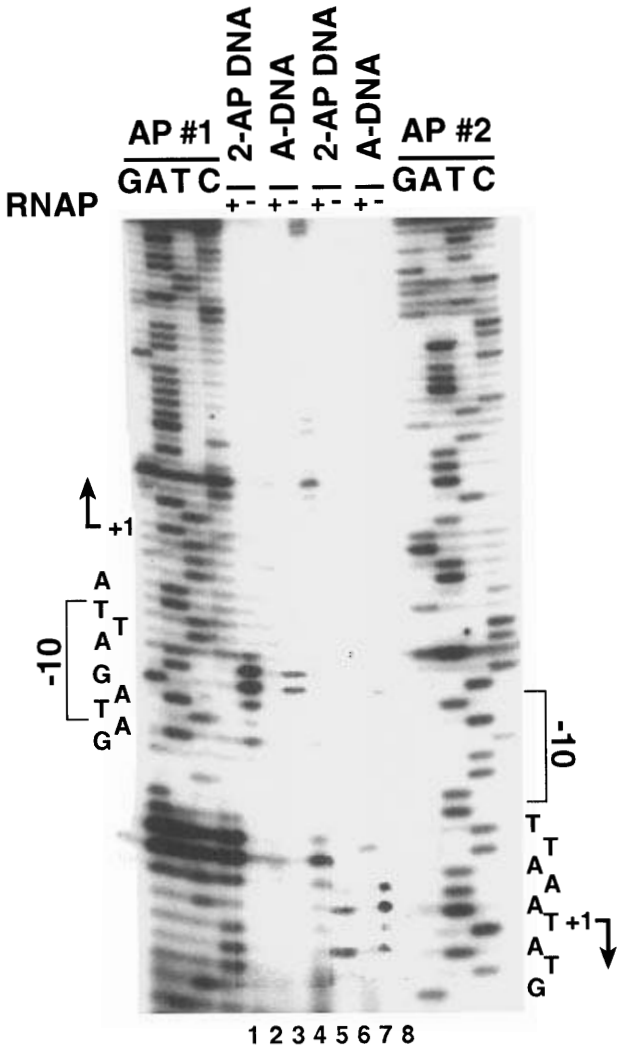

FIGURE 2: $KMnO_4$ probing of the region of **C** DNA opened by RNA polymerase. Both the sequences and the determination of the positions of modification of the promoter DNA employed 5′ ^{32}P end-labeled primers (see Figure 1), as indicated. Positions of RNA polymerase-induced modification of the nontemplate strand are deduced from the bands in lanes 1 (2-AP-substituted DNA) and 3 (unsubstituted) and of the template strand in lanes 5 (2-AP-substituted DNA) and 7 (unsubstituted). Lanes 2, 4, 6, and 8 respectively are identical to the above, but without added RNA polymerase.

was probed in both directions to more precisely define the edges of the strand-separated region. Comparison with the calibrating sequencing lanes on either side shows that, for variant **C** (and for **NC**, data not shown), the region of strand separation spans positions −9 (lanes 1−4) to +1 (lanes 5−8), regardless of whether the DNA contains 2-AP substitutions. Its location is as expected for promoters utilized by *E. coli*

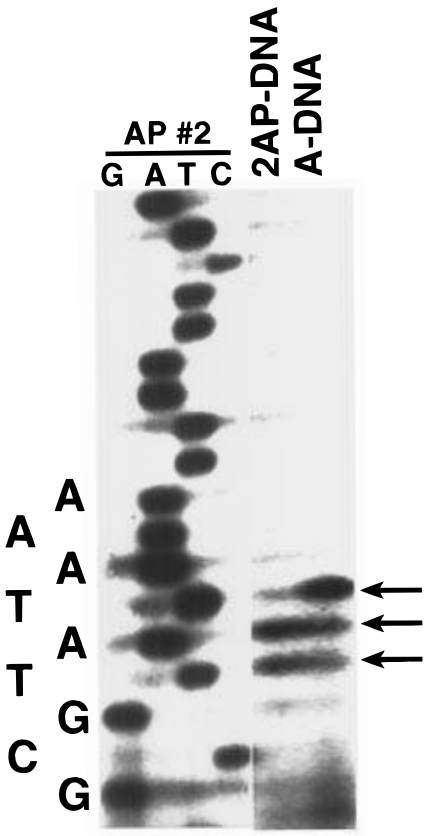

FIGURE 3: Determination of the start site for RNA synthesis. The identity of the 5′-most nucleotide of the RNA is revealed by that at the 5′ end of the DNA made using the RNA as template and primed by ^{32}P-labeled AP 2 (the DNAs are indicted by the arrows). Calibration of the gel is by DNA sequencing lanes using exactly the same primer. Sequencing and primer extension lanes are shown in different exposures.

RNA polymerase, including P_{RM} (*21*) of which **C** and **NC** are variants.

The results displayed in Figure 3 address the site of initiation of RNA synthesis. Runoff transcripts from **C** and **NC** are similar in size, but both are considerably longer than the 31 nucleotides expected, on the basis of the position of the start site (data not shown). This can be explained with the recently described phenomenon of "template switching", which involves the use of more than one template in the synthesis of a single runoff RNA product (*26*). Primer extension probing of the transcript from **C** clearly shows that the start site of transcription on templates without 2-AP is the same as has previously been observed with P_{RM} and its variants. Close inspection of the gels does unveil a difference in the initiation site on templates containing 2-AP; the preferred site is an "A" two base pairs downstream from the normally used A site. This may reflect a downstream extension of the melted bubble, possibly due to the lesser stability of the 2-AP-T base pairs (*5*).

Substitution of the base analog 2-AP affords the possibility of following RNA polymerase-induced strand separation of promoter DNA in real time, and thus of obtaining information about the rate of the strand-opening (or -closing) process. We first demonstrated the feasibility of using 2-AP-substituted DNA with the **NC** promoter. To determine whether the rate of strand opening was affected by the presence of 2-AP, we measured the kinetics of complex formation by the gel mobility shift assay (*21*); only open complexes have a long enough lifetime to be detected on the gel. We used four different versions of the **NC** DNA: containing no 2-AP, with the substitutions as indicated in Figure 1, and with 2-AP at the indicated positions on either one of the two strands. With each of the four templates, a first-order rate constant of $(7 \pm 2) \times 10^{-3}$ s^{-1} was obtained at 37 °C and an active RNA polymerase concentration of 50 nM (data not shown). This indicates that the presence of the 2-AP substitutions did not affect the rate of open complex formation despite the reduced stability of a 2-AP-T base pair. This observation is also consistent with the notion that strand separation is not the actual rate-limiting process.

Upon mixing a 2-AP-containing promoter fragment with RNA polymerase, an increase in the fluorescence signal (excitation at 310 nm, emission at 370 nm) is observed. The amplitude of this signal is approximately 2/3 of that of the sum of the individual signals obtained with the same concentration of the two 2-AP-bearing oligonucleotides (2 and 5 in Figure 1; data not shown). At equimolar ratios of RNA polymerase and 2-AP-substituted DNA, both contribute approximately equally to the observed signal in the 330−400 nm range. Subtraction of the signal due to the RNA polymerase results in a spectrum with a maximum at 370 nm (data not shown), characteristic of the emission spectrum of 2-AP (*9*). Control experiments demonstrated that no such signal was observed with DNA lacking 2-AP substitution, or with 2-AP-substituted DNA at a temperature of 5 °C, where promoter binding but no strand separation is expected to occur (*27*). From these results, as well as the known sensitivity of 2-AP to its environment (*8*), we conclude that the observed fluorescence signal reflects RNA polymerase-induced melting of DNA in the region containing the 2-AP substitutions.

Monitoring of the increase in fluorescence signal in a fluorometer after manual mixing of 10 nM 2-AP-substituted **NC** and 50 nM RNA polymerase yielded a k_{obs} of $(9 \pm 2) \times 10^{-3}$ s^{-1} at 37 °C, in good agreement with results obtained by the gel shift assays on this same promoter (see above). On the basis of these results, an apparent second-order rate constant of approximately 2×10^5 M^{-1} s^{-1} can be calculated ($k_a = k_{obs}$[RNAP]), if the reaction is modeled as a one-step process; this approximation is reasonable under the conditions of the experiment (see below). Data collected over a range of RNA polymerase concentrations revealed an apparently linear dependence of k_{obs} versus [RNAP] (thin line in Figure 4), but with a physically uninterpretable negative *y* axis intercept. We surmise that this behavior might be due to a lack of precision in our measurements, or that it is due to RNA polymerase binding to the ends of the fragment where it could interfere with interaction at the promoter. From the slope of the plot, k_a is determined to be $(2.4 \pm 0.4) \times 10^5$ M^{-1} s^{-1}, in good agreement with the observations above. If the fit is constrained to go through the origin (thick line in Figure 4), a shallower slope is obtained and $k_a = (1.6 \pm 0.2) \times 10^5$ M^{-1} s^{-1}; clearly, the values cluster about a $k_a = 2 \times 10^5$ M^{-1} s^{-1} for the **NC** promoter at 37 °C. Values for k_a determined at three different temperatures are collected in Table 1.

Open complex formation at the **C** promoter proceeded to completion in too short a time to monitor the process with a regular spectrofluorometer (data not shown). In Figure 5, the results are shown of an experiment in which the fluorescence signal at wavelengths of greater than 350 nm was followed after mixing of RNA polymerase and 2-AP-

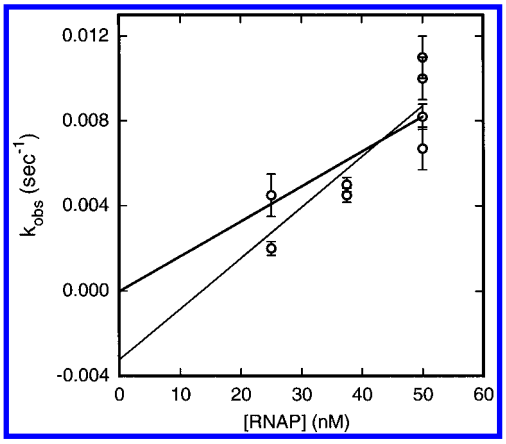

FIGURE 4: Dependence of observed rate constants (k_{obs}) for open complex formation at the **NC** promoter on the concentration of active RNA polymerase. Each independent determination was carried out by manual mixing and monitoring the development of the fluorescence signal in a cuvette. The DNA concentration was 10 nM. The thin line represents a least-squares fit through the data points [y intercept $= -(3.2 \pm 2.4) \times 10^{-3}$ s^{-1}) and the thick line a fit constrained to go through the origin.

substituted DNA **C** at 37 °C in a stopped flow apparatus. The data set is fit very well by a single-exponential, three-parameter equation, as described in Materials and Methods. The curve drawn through the data points is for an observed first-order rate constant, $k_{obs} = 0.69$ s^{-1}. The rate of complex formation at 37 °C on the **C** template as a function of the RNA polymerase concentration (at sufficient excess to maintain pseudo-first-order conditions) is shown in Figure 6. Each data point is the average of at least three separate determinations; error bars represent the standard deviation. The thicker line is the best fit of the data to a rectangular hyperbola, $k_{obs} = k_a$[RNAP]/(K_B[RNAP] + 1), representing a pre-equilibrium model, with an irreversible rate-limiting step for formation of an open complex (*28*):

Scheme 1

$$R + P \overset{K_B}{\rightleftharpoons} RP_c \overset{k_f}{\longrightarrow} RP_o$$

Here, K_B [$(0.8 \pm 0.3) \times 10^7$ M^{-1} for the data shown in Figure 6] is the equilibrium constant for the interaction of RNA polymerase (R) and promoter (P) which forms the closed complex (RP_c); k_f (2.5 ± 0.7 s^{-1}) is the first-order rate constant for the conversion of RP_c to the open complex, RP_o, and the second-order rate constant $k_a = K_B k_f$ (2×10^7 M^{-1} s^{-1}). In Table 1, we have collected kinetic data obtained with the **C** promoter at three different temperatures.

Also shown in Figure 6 is a linear fit with a correlation coefficient of 0.98 (thinner line); clearly, the error in the data does not allow an *a priori* evaluation of which of the two mechanisms affords a better analysis of the data. The simplest reaction scheme compatible with a linear dependence of k_{obs} on [RNAP] is shown below:

Scheme 2

$$R + P \underset{k_d}{\overset{k_a}{\rightleftharpoons}} RP_o$$

Here $k_{obs} = k_a$[RNAP] + k_d. The second-order rate constant for formation of open promoter complexes, k_a, as determined from the slope of the line is $(1.0 \pm 0.1) \times 10^7$ M^{-1} s^{-1}, in

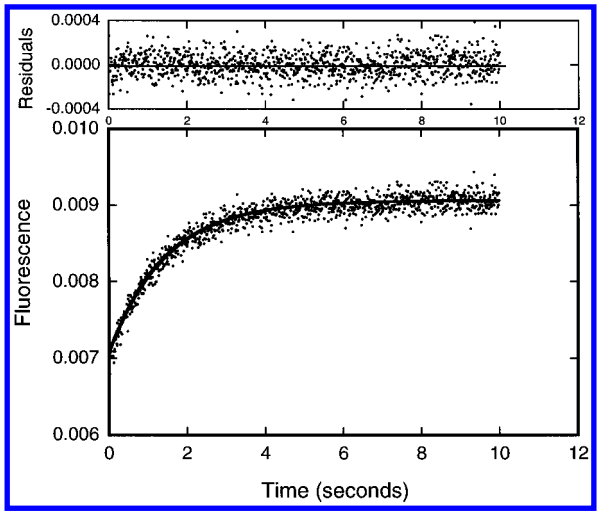

FIGURE 5: Time dependence of the increase in fluorescence upon mixing RNA polymerase (50 nM) with the **C** template (10 nM). The experiment was performed with a stopped flow thermostated at 37 °C. The fluorescence was excited at 315 nM and measured after passage through a 350 nm cut-on filter; 1000 data points were accumulated over the 10 s time course of the experiment. The line drawn through the data points is a one-exponential fit through the data points with a k_{obs} = of 0.69 ± 0.1 s^{-1}. Residuals are shown at the top.

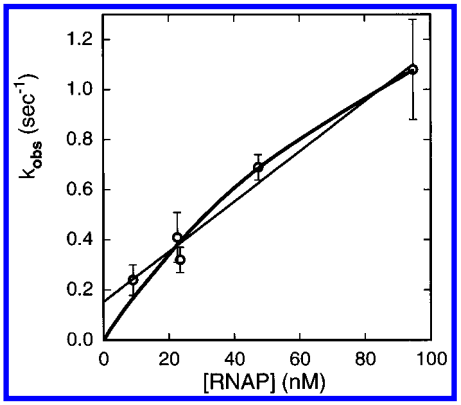

FIGURE 6: Dependence of observed rate constants (k_{obs}) for open complex formation at the **C** promoter on the concentration of active RNA polymerase. Error bars represent standard deviations determined from averaging multiple individual experiments at each concentration. The DNA concentration was 3 nM for the lowest concentration of RNA polymerase and one of the two determinations at the next lower concentration and 10 nM for the others. The linear (thin line) and hyperbolic fits (thicker line) are applicable to different models for the mechanism of formation of an open complex at the promoter (see the text).

good agreement with the value reported above. The intercept with the y axis (0.15 ± 0.04 s^{-1}; error estimate based on the least-squares fit) equals k_d; this would correspond to a half-life for the RP_o complex of about 5 s. Other models, more complex than that of Scheme 2, are also compatible with a non-zero intercept for a plot of k_{obs} versus [RNAP]. These will be considered in the Discussion.

From the slopes of plots of ln k_a (as determined from the product $K_B k_f$ for **C**) versus $1/T$ (not shown), the activation free energy for open complex formation was determined to be 15 kcal/mol in the 37−15 °C range. For the **NC** promoter, the temperature dependence of k_a (determined for the 37−25 °C range, see Table 1) also yielded an E_a of 15 kcal/mol. These values of E_a are lower than the 40 kcal/mol previously determined (*29*) for another variant of P_{RM} (same −35 region as **NC** and one nonconsensus bp in the

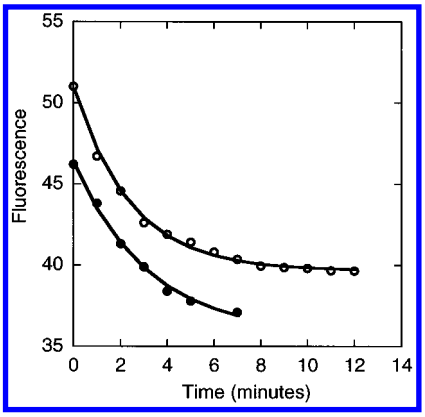

FIGURE 7: Time-dependent loss of the fluorescence signal upon addition of NTPs to preformed open complexes. Data collection was initiated immediately after addition of a mixture of ATP, CTP, GTP, and UTP (to a final concentration of 200 nM each) and heparin to open complexes formed at 37 °C. The data points are the averages of five and three individual experiments for **C** (open circles) and **NC** (closed circles), respectively. The curves represent one-exponential fits to these data, with rate constants of 0.42 ± 0.02 and 0.32 ± 0.02 min^{-1}, respectively, for the above two promoters.

−10) but similar to that for open complex formation at P_R (*30*). For the latter, E_a was reported to be 20 kcal/mol in the 37−25 °C range, and greater at lower temperatures. Our data set (see Table 1) is not extensive enough for either promoter to allow an assessment of whether the interaction of RNA polymerase with the constructs studied here also exhibits a temperature-dependent E_a.

As an assay for promoter clearance, we have also measured the rates of promoter annealing upon addition of nucleoside triphosphates to preformed open complexes for the constructs shown in Figure 1; see Figure 7. At the indicated zero time, all four nucleoside triphosphates and heparin were added to the complex and decay of the fluorescence signal was followed as a function of time. These data were fit adequately to a single exponential (see Materials and Methods). On the basis of multiple determinations, the calculated first-order rate constants of strand closing for the **C** and **NC** templates were (5.5 ± 2.5) × 10^{-3} and (7.7 ± 2.7) × 10^{-3} s^{-1}, respectively. These values are similar within the error of the assay and indicate a half-time of about 2 min for the observed loss of signal.

DISCUSSION

We have demonstrated the feasibility of spectroscopically monitoring open complex formation between *E. coli* RNA polymerase and short promoter-containing DNA fragments bearing 2-AP substitutions. For the **C** construct, no indication of interference was seen either from the substituted 2-AP nucleotides or from end binding by RNA polymerase, but possibly such interference comes into play for the weaker **NC** promoter. However, it should not be difficult to generate substantially longer synthetic fragments using the ligation scheme we employed, by increasing the sizes of the two outside pairs of oligodeoxynucleotides on either side of the 2-AP-labeled fragments. This would not only put the promoter region farther from the ends but also allow the inclusion of upstream control regions, such as UP elements (*11*), on the DNA templates.

With template **C**, we determined a value for k_a of (0.6 ± 0.1) × 10^7 M^{-1} s^{-1} at 25 °C. This can be compared with a $k_a = 10^7$ M^{-1} s^{-1} (*23*) obtained at the same temperature with this promoter contained on linear plasmid DNA in a buffer with the same monovalent cation concentration, but with pH 8.0 and 3 mM $MgCl_2$ instead of pH 7.6 and 10 mM as used here. Each of the two differences may have a sizable effect, but the two are predicted to be in opposite directions. We estimate the effects of the two differences approximately cancel on the basis of the pH and Mg^{2+} dependencies for the *equilibrium* binding constant for the interaction between *E. coli* RNA polymerase and T7 early promoters (*31*). Therefore, we conclude that having the promoter on an 86 bp fragment as compared to a 9 kb plasmid does not result in major deleterious effects. The experiments reported here show that k_a for **C** is about 100-fold greater than that for **NC**. We believe that this more accurately reflects the relative strengths of the two promoters than our previous estimate of 1000-fold (*23*), which was based on less extensive data.

With DNA **C** at 37 °C, we found k_a = 1−2 × 10^7 M^{-1} s^{-1} (depending on the method used to evaluate k_a). This result can be compared to the extensive filter binding studies of Record and co-workers on the interaction of *E. coli* RNA polymerase with the P_R promoter of bacteriophage λ (*30*). They found that $k_a = 2.7 \times 10^6$ at 37 °C in buffer at pH 8.0, containing 0.12 M KCl. Again using the results of Strauss et al. (*31*), we estimate that this value is about ten-fold lower than it would have been in our buffer (pH 7.6 and [KCl] = 0.1 M), thus establishing the **C** and P_R promoters as about equally fast in forming an open complex at 37 °C. Open complex formation at the **C** promoter is 1 order of magnitude slower than that at the rrnB promoter containing the UP element (*11*); $k_a = 10^8$ M^{-1} s^{-1} at pH 7.9 and 30 mM KCl, which may be near the diffusion limit. The difference between the rates may be as much a result of the differences in solution conditions as in the intrinsic strengths of the two promoters, but the fastest rate we observe is clearly at least 1 order of magnitude below the diffusion limit. As the difference cannot be explained on the basis of the size of the fragments used (see above), it has implications for the mechanism of open complex formation at the **C** promoter.

It is useful to review four constraints to any model for the process of open complex formation that can be used to interpret our data. First, the observed signal is due to an actual strand-opening process; we do not observe a signal when RNA polymerase and 2-AP-substituted DNA are mixed at 5 °C and no strand opening occurs. In addition, others have shown that the fluorescence of 2-AP-substituted DNA increases upon melting (*8*, *9*). Second, open complex formation is a multistep process. The slower than diffusion-limited reaction rate and the high energy of activation (see Results) both validate this contention. Others have addressed this question for various promoters (*28*, *30*, *32*, *33*) and concluded (see ref 1 for a recent review) that Scheme 2, while practically useful, is a substantial oversimplification of the process. Third, formation and dissociation of a closed complex are fast compared to the rate with which this complex is converted to complexes farther along the pathway of the reaction. For the P_{RM} promoter, this has been elegantly demonstrated by Hawley and McClure (*34*). Fourth, the final open complex is generally very stable, making its formation essentially an irreversible process. We have observed that addition of heparin to preformed complexes did not result in a reversion of the fluorescence enhancement over a 30 min period (data not shown). This indicates that dissociation

of the open complex is negligible over this time span, in agreement with our previous studies on variants of the P_{RM} promoter (*35*).

Scheme 1 is a minimal model which satisfies the criteria laid out above. It has been very useful for the interpretation of kinetic experiments carried out by manual mixing, and its validity has recently been reaffirmed by Record and co-workers (personal communication, 1997). The simplest modification of Scheme 2 which not only satisfies the above criteria but also preserves the reversibility of the formation of the complex responsible for the observed fluorescence signal is shown as Scheme 3:

Scheme 3

$$R + P \overset{K_B}{\rightleftharpoons} RP_c \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} RP_o^* \xrightarrow{k_3} RP_o$$

The first step is the rapid equilibrium between the closed complex and the free RNA polymerase and promoter DNA and the second the reversible formation of an open complex, RP_o^*, which would be responsible for the observed fluorescent signal, and the third would generate the stable, classically studied open complex RP_o without further contributing to the signal. This putative last step would have to occur on a time scale of tens of seconds to a few minutes in order to be both slower than dissociation of RP_o^* (the back-reaction) and fast enough to have escaped detection in manual-mixing experiments. Scheme 3, just as Scheme 1, would result in a hyperbolic fit; the data shown in Figure 6 would be in the quasi-linear range of the hyperbola (K_B[RNAP] $\ll$ 1), and a substantially larger [RNAP] would be required to approach a plateau value for k_{obs}.

The crucial difference between Schemes 1 and 3 is the *y* axis intercept. An intercept of 0 would indicate that Scheme 1 was the simplest interpretation of the data, but a non-zero intercept would require the introduction of the unstable open complex intermediate, RP_o^*, as shown in Scheme 3. While the data in Figure 6 appear to support an interpretation compatible with Scheme 3, there are insufficient data points in the region of low [RNAP] to allow a clear differentiation between the two models. We have probed the stability of the earliest detectable open complex by a manual quench experiment. Within 15 s of mixing RNA polymerase (100 nM active enzyme) and construct **C** (25 nM), heparin was added to inactivate free RNA polymerase and thus enable an investigation of the stability of the complex formed during the prior short incubation. Analysis of the samples by gel electrophoresis demonstrated the formation of a gel-shifted complex within 10 s, in agreement with the data presented in Figure 5. Within 15 s of heparin addition, less than 15% dissociation of this species was observed, and increased incubation times did not result in further dissociation (J. J. Sullivan and P. L. deHaseth, unpublished). On the basis of the model represented in Scheme 3, this would indicate that at least 85% of the complex had already been converted to the stable form RP_o in 15 s, with an estimated rate constant of 0.13 s^{-1} and a relaxation time of 8 s. This would be similar to the rate of back-reaction of the postulated RP_o*, which is inconsistent with the assumption on which Scheme 3 is based, namely that the formation of RP_o* would be reversible and the back-reaction thus substantially faster than the formation of RP_o. Thus, we conclude that it is unlikely that the fluorescent signal we detect reflects the formation of an unstable open complex intermediate at the **C** promoter.

We have also demonstrated the feasibility of using 2-AP as a probe for monitoring promoter clearance upon the addition of NTPs to a preformed open complex. In studies of the effect of promoter strength on the rates of open complex formation and promoter clearance, an inverse relationship was found between the rates of these two processes (*36*, *37*). Possibly, it is more difficult for RNA polymerase to leave stronger promoters with which it initially established more favorable contacts. As our data (see Table 1) indicate a 100-fold greater rate constant for open complex formation at **C** than at **NC**, we expected that after addition of NTPs the **C** promoter would close at a slower rate than **NC**. However, no significant difference was observed between the two promoters. For two promoters differing by only one base pair (P_{ant} and its consensus variant), Smith and Sauer (*37*) observed that P_{ant} was cleared faster. The promoters used by these workers differed in the sequence of their −10 region, while ours have different −35 regions. It may be that the −10 region is a more important determinant of promoter clearance and that an entirely consensus −10 region specifies an especially slow rate for this process (*36*, *37*). The rate constants for promoter clearance that can be calculated from our data are in surprisingly good agreement with those determined for P_{bla} (*38*) and P_{ant} (*37*). With measurement of rates for single-round runoff RNA synthesis from P_{RM} promoter variants, very similar results have again been obtained (F. Campbell, D. R. Setzer, and P. L. deHaseth, unpublished). Apparently, the rate of promoter clearance is slow compared to the rates of open complex formation or elongation. Thus, *in vitro*, promoter clearance is the limiting step in the initiation of RNA synthesis from the P_{RM}, P_{ant}, RNAI, and bla promoters. For P_{ant} and its consensus variant, it was shown that Arc protein was able to accelerate this process (*37*). GreA and GreB have also been observed to facilitate promoter clearance (*39*).

While our experiments were in progress, three reports appeared on the use of short, 2-AP-substituted, synthetic promoters for monitoring open complex formation with T7 RNA polymerase (*14*−*16*). Beechem and co-workers described the use of 2-AP-substituted promoters to monitor transient *E. coli* RNA polymerase-induced strand separation at a synthetic promoter (*13*) during the process of RNA synthesis. In addition, these workers used two other probes to detect RNA polymerase binding, independent of strand opening, as well as to monitor phosphodiester bond formation during the course of RNA synthesis. Our work and the latter study are complementary. We have demonstrated that 2-AP substitution does not greatly affect the region of strand separation or the position of the start site and that end effects are not a problem in the study of strong promoters. Thus, our results validate the use of this base analog on relatively short synthetic DNA fragments. While Dunkak et al. (13) established the feasibility of using 2-AP to detect the strand separation process, the kinetic studies presented by these workers were obtained by monitoring fluorescence anisotropy of a probe attached at one end of an 80 base pair synthetic fragment bearing a consensus promoter sequence. The time course of RNA polymerase binding at 37 °C is similar to that shown in Figure 5, and association rate constants of 1.5 $\times$ 10^7 to 7.4 $\times$ 10^7 M^{-1} s^{-1} are obtained from data initially

subjected to, respectively, one- and two-exponential fits; these values are within 1 order of magnitude of that reported here for **C** at 37 °C. An as yet unexplained feature of the work of Dunkak et al. (*13*) is the slow appearance of both the 2-AP fluorescence signal and that due to RNA synthesis (estimated half-times of at least 10 min). It does not appear that these observations are related to the slow re-annealing we observe upon initiation of RNA synthesis (see Figure 7).

ACKNOWLEDGMENT

We thank Dr. T. Lohman for the use of his stopped flow apparatus, Dr. J. Leis for the use of the fluorometer, Dr. S. Patel for conducting preliminary stopped flow experiments, K. Sullivan for doing gel shift experiments, and Drs. V. Anderson, T. Lohman, and M. T. Record, Jr., for helpful discussions.

REFERENCES

1. Record, M. T., Jr., Reznikoff, W. S., Craig, M. L., McQuade, K. L., and Schlax, P. J. (1996) in *Escherichia coli and Salmonella, cellular and molecular biology* Neidhardt, F. C., Ed.) pp 792−820, ASM Press, Washington, DC.
2. Sowers, L. C., Fazakerley, G. V., Eritja, R., Kaplan, B. E., and Goodman, M. F. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5434−5438.
3. Nordlund, T. M., Anderson, S., Nilsson, L., Rigler, R., Graslund, A., and McLaughlin, L. W. (1989) *Biochemistry 28*, 9095−9103.
4. Wu, P., and Nordlund, T. M. (1990) *Biochemistry 29*, 6508−6514.
5. Law, S. M., Eritja, R., Goodman, M. F., and Breslauer, K. J. (1996) *Biochemistry 35*, 12329−12337.
6. McLaughlin, L. W., Benseler, F., Graeser, E., Piel, N., and Scholtissek, S. (1987) *Biochemistry 26*, 7238−7245.
7. Ward, D. C., Reich, E., and Stryer, L. (1969) *J. Biol. Chem. 244*, 1228−1237.
8. Guest, C. R., Hochstrasser, R. A., Sowers, L. C., and Millar, D. P. (1991) *Biochemistry 30*, 3271−3279.
9. Raney, K. D., Sowers, L. C., Millar, D. P., and Benkovic, S. J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 6644−6648.
10. McClure, W. R. (1985) *Annu. Rev. Biochem. 54*, 171−204.
11. Rao, L., Ross, W., Appleman, J. A., Gaal, T., Leirmo, S., Schlax, P. J., Record, M. T., Jr., and Gourse, R. L. (1994) *J. Mol. Biol. 235*, 1421−1435.
12. Tang, Y., Murakami, K., Ishihama, A., and deHaseth, P. L. (1996) *J. Bacteriol. 178*, 6945−6951.
13. Dunkak, K. S., Otto, M. R., and Beechem, J. M. (1996) *Anal. Biochem. 343*, 234−244.
14. Jia, Y., Kumar, A., and Patel, S. S. (1996) *J. Biol. Chem. 271*, 30451−30458.
15. Sastry, S. S., and Ross, B. M. (1996) *Biochemistry 35*, 15715−15725.
16. Ujvari, A., and Martin, C. T. (1996) *Biochemistry 35*, 14574−14582.
17. Burgess, R. R., and Jendrisak, J. J. (1975) *Biochemistry 14*, 4634−4638.
18. Gonzales, N., Wiggs, J., and Chamberlin, M. J. (1977) *Arch. Biochem. Biophys. 182*, 404−408.
19. Fujimoto, J., Nuesca, Z., Mazurek, M., and Sowers, L. C. (1996) *Nucleic Acids Res. 24*, 754−759.
20. deHaseth, P. L., Goldman, R. A., Cech, C. L., and Caruthers, M. H. (1983) *Nucleic Acids Res. 11*, 773−787.
21. Mita, B. C., Tang, Y., and deHaseth, P. L. (1995) *J. Biol. Chem. 270*, 30428−30433.
22. Aiyar, S. E., Helmann, J. D., and deHaseth, P. L. (1994) *J. Biol. Chem. 269*, 13179−13184.
23. Szoke, P. A., Allen, T. A., and deHaseth, P. L. (1987) *Biochemistry 26*, 6188−6194.
24. Melançon, P., Burgess, R. R., and Record, M. T., Jr. (1983) *Biochemistry 22*, 5169−5176.
25. Sasse-Dwight, S., and Gralla, J. D. (1989) *J. Biol. Chem. 264*, 8074−8081.
26. Nudler, E., Avetissova, E., Markovtsov, V., and Goldfarb, A. (1996) *Science 273*, 211−217.
27. Kovacic, R. T. (1987) *J. Biol. Chem. 262*, 13654−13661.
28. McClure, W. R. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5634−5638.
29. Hershberger, P. A., Mita, B. C., Tripatara, A., and deHaseth, P. L. (1993) *J. Biol. Chem. 268*, 8943−8948.
30. Roe, J. H., Burgess, R. R., and Record, M. T. (1985) *J. Mol. Biol. 184*, 441−453.
31. Strauss, H. S., Burgess, R. R., and Record, M. T., Jr. (1980) *Biochemistry 19*, 3504−3515.
32. Buc, H., and McClure, W. R. (1985) *Biochemistry 24*, 2712−2723.
33. Suh, W. C., Leirmo, S., and Record, M. T., Jr. (1992) *Biochemistry 31*, 7815−7825.
34. Hawley, D., and McClure, W. R. (1982) *J. Mol. Biol. 157*, 493−525.
35. Auble, D. T. (1987) Ph.D. Thesis, Case Western Reserve University, Cleveland, OH.
36. Ellinger, T., Behnke, D., Bujard, H., and Gralla, J. D. (1994) *J. Mol. Biol. 239*, 455−465.
37. Smith, T. L., and Sauer, R. T. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 8868−8872.
38. Schmitt, B., and Reiss, C. (1995) *Biochem. J. 306*, 123−128.
39. Hsu, L. M., Vo, N. M., and Chamberlin, M. C. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 11588−11592.

BI970363K