Kinetic Analysis of T7 RNA Polymerase Transcription Initiation from Promoters Containing Single-Stranded Regions†

Maribeth Maslak and Craig T. Martin*

Department of Chemistry, University of Massachusetts, Amherst, Massachusetts 01003

Received September 2, 1992; Revised Manuscript Received February 11, 1993

ABSTRACT: T7 RNA polymerase is highly specific for the initiation of transcription from a relatively small consensus promoter sequence. Previous footprinting studies suggested that the enzyme binds specifically to a fully closed duplex form of the promoter, recognizing functional groups along one face of the helix [Muller, D. K., Martin, C. T., & Coleman, J. E. (1989) *Biochemistry 28*, 3306–3313]. Steady-state kinetic analysis of oligonucleotide-based promoters shows that removal of the nontemplate strand completely within the message region of the DNA (positions +1 through +5) results in no change in binding (as reflected in the parameter K_m) and a 2-fold increase in kinetics (as reflected in k_{cat}). Further deletion of the nontemplate strand as far upstream as position −4 has no effect on binding, and although deletion upstream through position −6 weakens binding, specific initiation continues at a high rate. The temperature dependence of the initiation kinetics shows a single apparent activation energy of ≈26 kcal/mol for the fully duplex promoter. Similar measurements on the promoter lacking the nontemplate strand in the message region show that less than 10% of this barrier is related to melting of the downstream region of the promoter. These results lead us to revise the previous model for recognition to include specific binding to a form of the promoter which is duplex upstream of about position −6 and melted downstream through the start site. Within the melted region, the polymerase interacts significantly only with the template strand of the promoter DNA.

T7 RNA polymerase is the best studied of a family of highly specific, single-subunit RNA polymerases (Chamberlin & Ryan, 1982). Its simplicity and apparent lack of regulation make T7 RNA polymerase an ideal model system in which to understand fundamental aspects of promoter recognition and the initiation of transcription. Previous kinetic studies have demonstrated that a small 17 base pair consensus sequence is sufficient for maximal binding and kinetics (Martin & Coleman, 1987). This kinetic approach has allowed a variety of studies which are not possible with phage or plasmid-based DNA templates.

Footprinting studies of the T7 RNA polymerase–promoter complex have generally confirmed the small size of the promoter (Ikeda & Richardson, 1986; Basu & Maitra, 1986; Gunderson et al., 1987; Chapman et al., 1988), and studies with Fe(II)EDTA·H_2O_2 as the footprinting reagent have provided further detail, leading to a simple model for promoter recognition (Muller et al., 1989). The four distinct regions of the promoter DNA that are protected from attack by hydroxyl radical in the enzyme-bound complex are compared in Figure 1. Projection of these regions onto a standard B-form DNA helix demonstrates that the spacing and phasing of these regions are highly suggestive of an enzyme bound to one face of fully duplex DNA. Indeed, similar patterns, with similar interpretations, have been observed for the interaction of the RNA polymerase from *Escherichia coli* with its much larger promoter (Metzger et al., 1989; Schickor et al., 1990; Mecsas et al., 1991).

In contrast to the apparent simplicity of the footprinting results, other studies of promoter binding by T7 RNA polymerase have indicated that a region of the nontemplate strand of the promoter near the start site (positions −6 to +2) shows enhanced reactivity to a single-strand specific endonuclease (Strothkamp et al., 1980; Osterman & Coleman, 1981; Muller et al., 1989). In addition, it has been demonstrated that T7 RNA polymerase can synthesize normal transcripts from a promoter lacking the nontemplate strand downstream of and including approximately position −5 (Milligan et al., 1987). Finally, recent studies of promoter recognition in which thymidine is selectively replaced by deoxyuridine at specific positions within the promoter reveal that thymine methyl groups on the template strand at positions −1 and −3, which are predicted to lie on the back face of the duplex in the simple model, are nevertheless involved in recognition (Maslak et al., 1993). These data would not appear to be consistent with the proposed requirement that the enzyme recognize fully duplex DNA throughout the entire promoter sequence.

An oligonucleotide-based steady-state kinetic assay for transcription allows quantitative dissection of specific contributions to the initiation of transcription (Martin & Coleman, 1987; Maslak et al., 1993). In particular, the kinetics of synthesis of a very short (five base) RNA transcript follows the simple steady-state model shown below.

$$\text{Enz} + \text{DNA} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{Enz–DNA} \xrightarrow[\text{NTP's}]{k_{cat}} \text{Enz} + \text{DNA} + \text{RNA} \qquad K_m = \frac{k_{-1} + k_{cat}}{k_1} \quad (1)$$

From fits of velocity as a function of enzyme and DNA concentration, one can readily measure K_m and k_{cat} for initiation. To the extent that k_{-1} is substantially larger than k_{cat} (≈28 min^{-1}), K_m becomes an approximation to a dissociation constant for the enzyme–promoter complex. Thus, to a first approximation, this analysis allows one to differentiate effects on binding (K_m) and on initiation kinetics (k_{cat}).

The current study uses the kinetic approach outlined above to determine precisely the structural requirements for the promoter-specific initiation of transcription by T7 RNA polymerase. Analysis of kinetic parameters associated with

† Supported by Grant NP-722 from the American Cancer Society.

0006-2960/93/0432-4281$04.00/0

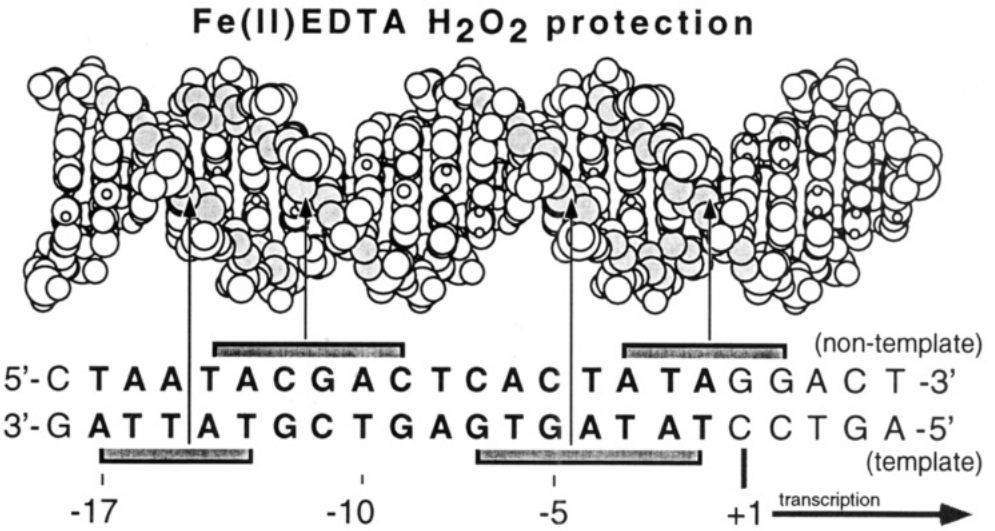

FIGURE 1: T7 consensus promoter sequence. Regions of the promoter protected from hydroxyl radical attack in the polymerase bound complex are filled in gray.

a variety of templates in which the nontemplate strand is partially deleted suggests that promoter recognition involves a form of the promoter DNA which is melted near the start site for transcription.

MATERIALS AND METHODS

T7 RNA Polymerase. T7 RNA polymerase was prepared from *E. coli* strain BL21 carrying the overproducing plasmid pAR1219, which has the polymerase gene cloned under inducible control of the *lac* UV5 promoter (Davanloo et al., 1984). T7 RNA polymerase was purified as previously described (King et al., 1986) by fractionation with Polymin P (less than 1.25%) and ammonium sulfate, followed by chromatography on Trisacryl SP-M (IBF Biotechnics Inc.), TSK-GEL Toyopearl CM-650 M (Supelco), and TSK-GEL Toyopearl DEAE-650 M (Supelco). A molar extinction of $\epsilon_{280} = 1.4 \times 10^5$ M^{-1} was used to determine enzyme concentrations (King et al., 1986). The purity of the enzyme was checked by SDS–polyacrylamide gel electrophoresis.

Kinetic Assay of Transcription Initiation. Oligonucleotides were synthesized and purified as described in Schick and Martin (1993). Kinetic assays of transcription (Martin & Coleman, 1987) were carried out in a total volume of 20 μL containing 0.005–0.16 μM DNA in 30 mM HEPES, pH 7.8, 15 mM magnesium acetate, 100 mM potassium glutamate, 0.25 mM EDTA,[1] 1 mM DTT, 0.1 mg/mL *N*,*N*-dimethylated casein (Sigma), 0.05% TWEEN-20 (Calbiochem, protein grade), 0.8 mM GTP, and 0.4 mM ATP, as described in Maslak et al. (1993).

Analysis of Transcription Products. Transcription products from each promoter were analyzed (data not shown) by denaturing polyacrylamide gel electrophoresis (7 M urea, 12–18% polyacrylamide). In all cases the correct five base RNA was the predominant product. In some cases, four base RNA products were also produced in low amounts and were included as total product formed when analyzed in the kinetic study.

RESULTS

Representative sets of data for specific promoter constructs are compared in Figure 2. The dependence of the velocity of production of a five base RNA transcript follows a dependence on total enzyme and DNA concentration as predicted by the exact solution to the simple steady-state kinetic formalism depicted in eq 1. For this study, promoter DNA constructs were prepared as shown in Figure 3. The same template strand (T7 promoter consensus sequence from positions –17 to –1, followed by DNA from +1 to +5 coding for the message GGGAA) was paired with a variety of complementary nontemplate strands to produce a series of constructs in which the nontemplate strand is successively removed near the start site for transcription. The top construct (+5) is the native template used in kinetic assays and is fully duplex. Comparison in Figure 3 of the K_m and k_{cat} parameters from this template with those from the truncated templates directly tests the role of the nontemplate strand near the start site for transcription.

Effects on K_m. Under conditions of our current assay, the best fit K_m value for transcription from the fully duplex consensus promoter sequence is approximately 2 nM. Transcription from a template double-stranded in the consensus promoter sequence from position –17 to –1, but single stranded throughout the message region (construct –1), yields a K_m value (1.9 nM) unchanged from that of the native template. In fact, removal of the nontemplate strand through position –4 (constructs –1, –4, and –5) results in no large change in K_m. Removal of bases in the nontemplate strand upstream to position –5 (construct –6) results in an approximately 5-fold increase in K_m (to 12 nM), while removal of the next base (construct –7) produces another 5-fold increase in K_m (to 47 nM). Continuing base removal (construct –8) yields a promoter so weakened that K_m can no longer be fit accurately by nonlinear regression analysis but is consistent with $K_m >$ 300 nM. For all constructs up to this point, the correct five base runoff transcript is the predominant RNA product (gel electrophoresis data not shown). Finally, construct –9 results in a promoter so weakened that no significant RNA products are produced. To the extent that K_m approximates a dissociation constant (K_d) for the enzyme–promoter complex, these results indicate that duplex DNA downstream from a region near position –4 to –5 is not required for optimum promoter binding. However, duplex DNA upstream from a region near position –8 is essential for even weak promoter binding.

Effects on k_{cat}. Under our current conditions, the best fit value of k_{cat} for fully duplex DNA is about 28 min^{-1}. Deletion of the nontemplate strand from positions +1 to +5 (construct

[1] Abbreviations: EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol.

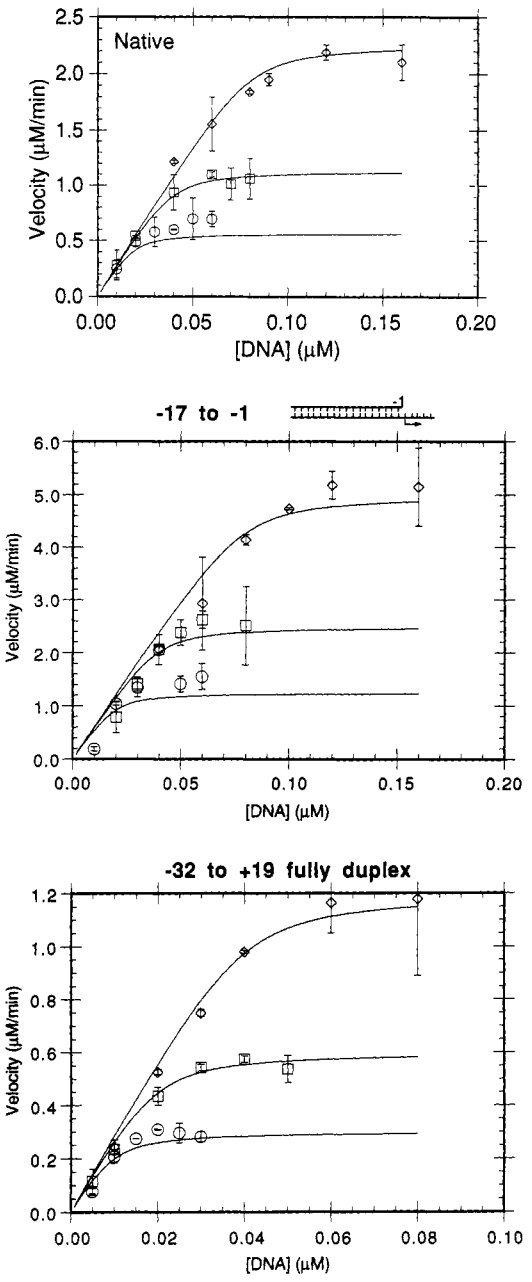

FIGURE 2: Representative steady-state kinetic fits for the initiation of transcription. The upper and middle panels present, respectively, a complete set of data for the fully duplex promoter (K_m = 2.0 nM; k_{cat} = 28.3 min^{-1}) and for the promoter which is single stranded within the message region (K_m = 1.9 nM; k_{cat} = 62.3 min^{-1}). The enzyme concentrations are (○) 0.02, (□) 0.04, and (◇) 0.08 μM. The lower panel presents data for a fully duplex promoter extending from position −32 to +19 (K_m = 1.7 nM; k_{cat} = 30 min^{-1}). The enzyme concentrations are (○) 0.01, (□) 0.02, and (◇) 0.04 μM. Other details of the experiment are described in the text.

−1) should remove the presumably fundamental requirement that the polymerase melt that part of the DNA directly involved in template-directed RNA synthesis (positions +1 and forward). The kinetic parameter k_{cat} doubles on this modified template, suggesting that melting of this portion of the DNA is a small barrier to the rate-limiting step(s) reflected in k_{cat}. Removal of three more bases in the nontemplate strand (construct −4) produces only a 15% reduction from this elevated k_{cat}. However, further removal of the nontemplate strand upstream into the consensus promoter yields a slow decrease in k_{cat}, such that the construct extending downstream only to position −7 has a k_{cat} similar to that of the native template. It appears that although removing the requirement for DNA

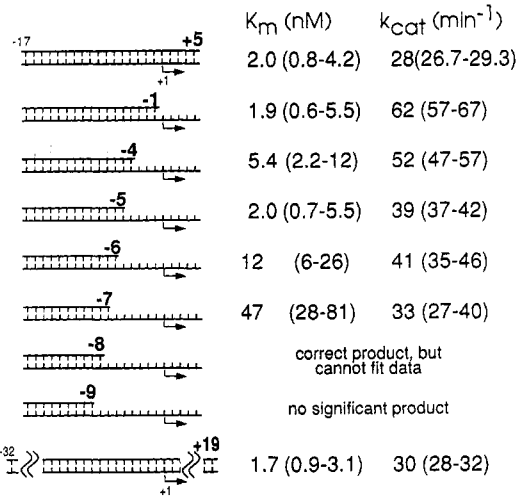

FIGURE 3: Comparison of best-fit steady-state kinetic parameters for fully duplex (−17 to +5) promoter and for similar promoters lacking portions of the nontemplate strand at the downstream end. The bottom construct represents a fully duplex promoter extending from −32 to +19, but encoding the same five base RNA transcript under current reaction conditions. Ranges in the parameters reflect the 65% joint confidence interval for each fit.

Table I: Temperature Dependence of Steady-State Kinetic Parameters from Duplex and Partially Single-Stranded Promoters

T (°C)	K_m (nM)	k_{cat} (min^{-1})
	Full Duplex	
22	3.0 (0.6–12)	3.3 (2.9–3.8)
27	5.2 (1.6–15)	7.7 (6.7–8.6)
32	6.0 (2.9–12)	16.5 (15.0–18.0)
37	2.0 (0.8–4.2)	28.3 (26.7–29.6)
42	3.0 (1.5–5.6)	55.2 (52.7–57.5)
	−17 to −1 (ssDNA Template)	
22	1.1 (0.4–3.1)	8.0 (7.6–8.3)
27	12.1 (5.6–26)	18.9 (16.0–21.7)
32	1.7 (0.5–4.6)	35.3 (32.5–37.8)
37	1.9 (0.6–5.5)	62.3 (57.1–66.9)
42	1.1 (0.2–3.7)	76.4 (70.7–80.9)

melting slightly increases k_{cat}, extensive removal of the nontemplate strand within the promoter to position −6 results in a similarly small decrease in k_{cat}.

These results indicate that melting of the DNA near the start site presents only a slight barrier to the initiation of transcription. In order to ensure that the original fully duplex construct (+5) is not already partially premelted due to thermal fraying of the ends of the DNA, we compared initiation kinetics from a template with an extended duplex message region. This construct, consisting of fully double-stranded DNA from position −32 to +19, codes for RNA with the initial primary sequence GGGAAC. Thus transcription in the presence of only the nucleotides GTP and ATP results in an abortive transcript of the same length and composition as in the above studies (confirmed by gel electrophoresis, data not shown). The kinetic parameters associated with this promoter template are indistinguishable from those of the shorter "native" 22 base template, demonstrating that, in the short templates, k_{cat} is not artificially high due to partial fraying of the downstream end of the duplex oligonucleotide.

Temperature Dependence of K_m and k_{cat}. In order to more fully understand the nature of the changes in k_{cat}, we compare in Table I the temperature dependence of both k_{cat} and K_m in transcription from the fully duplex native oligonucleotide promoter with the kinetic parameters from the construct lacking the nontemplate strand within the message region

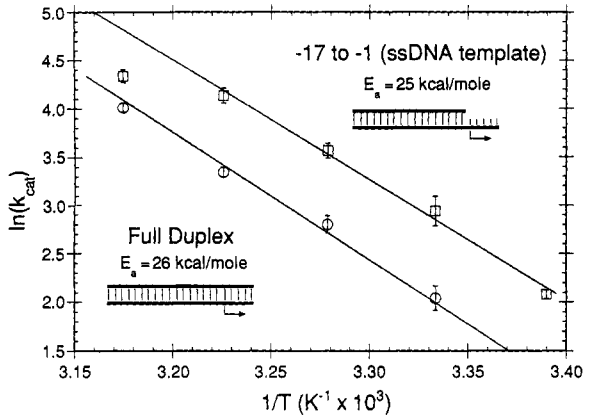

FIGURE 4: Temperature dependence of the kinetic parameter k_{cat} for transcription from the fully duplex promoter and from the promoter which is single stranded within the message region. The apparent activation energy, E_a, is derived from the best fit slope to data in the 22–37 °C temperature range.

(construct −1 in Figure 3). Changes in K_m are mostly small and show no obvious trends for either the fully duplex or the partially single-stranded template. The observed increase in K_m associated with the partially single-stranded template at 27 °C appears to be significant and is reproducible, but it is poorly understood at this time and is under further investigation. The parameter k_{cat}, however, shows a strong temperature dependence, and the Arrhenius analysis shown in Figure 4 suggests a simple kinetic rate constant. The apparent activation energy for k_{cat} on the fully duplex template is 26 ± 3 kcal/mol, while that associated with the template single-stranded in the message region is the same within error, 25 ± 3 kcal/mol (the error in each velocity is derived from the *t* distribution 80% confidence interval of the fitted slope for the data in the temperature range 22–37 °C). Similar results are found for transcription from the promoter construct single-stranded from position −5 downstream (preliminary data not shown). In this analysis, only the measurements from 22 to 37 °C were included, as preliminary results of the temperature dependence of the enzyme stability suggest that unfolding of the protein begins near 42 °C. From these apparent activation energies, we can conclude that promoter melting contributes less than 10% to the overall apparent rate-limiting step in initiation. This result reinforces the conclusion that although DNA melting does contribute to the rate-limiting step(s) in initiation, for the consensus promoter sequence it is only a minor contributor.

DISCUSSION

The pattern of protection from reaction with hydroxyl radical observed for T7 RNA polymerase bound to its promoter is highly suggestive of a model for promoter binding in which the enzyme interacts with one face of fully closed duplex DNA throughout the entire promoter from position −17 to at least position +1 (Muller et al., 1989). McAllister and Carter (1980) had previously proposed a three-domain promoter model, comprising melting, recognition, and initiation domains. As a result of in vitro studies of base pair substitutions in the T7 promoter, Chapman and Burgess (1987) later proposed a model for promoter function involving two domains. The first domain, encompassing positions −16 through −5, was proposed to specify binding determinants, while the second domain, encompassing positions −4 through +5, was proposed to encode "initiation" determinants. Indeed, footprinting studies with a single strand specific endonuclease argue that the latter domain is at least transiently single-stranded in the bound complex (Strothkamp et al., 1980; Osterman & Coleman, 1981), and functional studies have shown that the nontemplate strand near the start site is not required for activity (Milligan et al., 1987).

It is clear that at some point in the initiation process the DNA must melt to expose the template strand for Watson–Crick pairing with incoming ribonucleoside triphosphates, and it is likely that the overall recognition of the promoter might involve kinetic as well as thermodynamic components (McClure, 1985; Buc & McClure, 1985). In the current study, we not only test the various domain models for promoter binding but also directly probe the contribution of specific strands of the promoter to both binding and catalysis, as reflected in the kinetic parameters K_m and k_{cat}.

Evidence for Recognition of an Open Complex. To the extent that preequilibrium conditions apply for the intermediate state or states in the initiation of transcription, the steady-state kinetic parameter K_m approximates the dissociation constant, K_d, for the kinetically competent enzyme–DNA complex. In any case, changes in binding should be reflected in the value of K_m, and we have observed in this and other studies that K_m and k_{cat} can vary independently (Maslak et al., 1993; Schick & Martin, 1993). The result that K_m does not increase substantially as the nontemplate strand of the promoter is successively removed up to and including position −4 indicates either that the protein makes no energetically significant contacts with this part of the nontemplate strand of the promoter or that contacts with the nontemplate strand do occur, but positive and negative contributions are closely balanced and cancel. The exposure in the enzyme–promoter complex of the nontemplate strand to a single strand specific endonuclease (Strothkamp et al., 1980) is consistent with a lack of direct contacts between polymerase and the nontemplate strand, while the uniformity of the hydroxyl radical footprint (Muller et al., 1989) is consistent with the presence of direct contacts along the nontemplate strand (and therefore the energetics must balance). It has recently been shown that DNA triplex formation in the message region of the DNA (downstream of position +1) inhibits binding of the enzyme to promoter (Maher, 1992). It is possible that triplex formation interferes with melting of the DNA near the initiation site and therefore with promoter specific binding.

In any case, comparisons of K_m values for the truncated constructs indicate that protein–DNA contacts on the nontemplate strand *are* involved in binding upstream of and including position −5. The increase in K_m values for successive constructs removing single bases from the nontemplate strand in the region from −5 to −7 is approximately 5-fold, suggesting a net contribution of about 1 kcal/mol for each nucleotide of the nontemplate strand in this region. This can be compared to results of various functional group substitutions within the DNA, in which removing a single amino or methyl group can result in equal or larger changes in K_m (Maslak et al., 1993; Schick & Martin, 1993). It is likely that, in this "transition" region, contacts with the template strand continue to predominate. Indeed, recent studies have shown that the removal of a single thymine methyl from the template strand at position −6 results in a significant decrease in binding (Maslak et al., 1993).

The Kinetics of Initiation. For the fully duplex promoter construct, the temperature dependence of k_{cat} follows the Arrhenius formalism very well over the temperature range 22–37 °C. This behavior is consistent with a single apparent activation energy for the initiation of transcription (≈26 kcal/mol). Although the range of temperatures is limited, this

result suggests that the kinetic barrier to initiation is dominated by a single process.

Comparison of k_{cat} values for hybrid promoters in which a portion of the nontemplate strand is effectively removed provides some understanding of the nature of the rate-limiting steps in the initiation of transcription. If promoter melting is the major barrier to initiation, then removal of the nontemplate strand in the message region should lead to the loss of this barrier and to a corresponding increase in k_{cat}. Comparison in Figure 3 of steady-state kinetic parameters for various truncated templates reveals that the rate of initiation (as reflected in k_{cat}) increases 2-fold following the complete removal of the nontemplate strand in (only) the message region (+1 to +5). This suggests that melting of the duplex DNA downstream of and including the start site for transcription does present some barrier to initiation. However, for the consensus promoter sequence this barrier is small; the difference in apparent activation energies for the rate-limiting step is less than 3 kcal/mol, compared to an overall apparent activation energy of approximately 26 kcal/mol. It seems likely that the consensus promoter sequence has been optimized to facilitate promoter melting, such that melting of the DNA does not present the major kinetic (or thermodynamic) barrier to transcription.

Further removal of nontemplate strand bases upstream into the promoter slightly decreases k_{cat}. The very dramatic change in local environment accompanying the removal of a large stretch of the nontemplate strand within the promoter most likely has some effect on the rate-limiting step(s) in initiation. The small size of this effect, combined with the fact that transcription begins precisely at the correct position even after very severe disruptions of the promoter, underscores the lack of involvement of the nontemplate strand.

That promoter melting does not represent the major rate-determining step in initiation is in agreement with kinetic studies of transcription by the multisubunit *E. coli* RNA polymerase, in which a step following promoter melting is shown to be rate limiting (Buc & McClure, 1985). It remains to be seen at what point (below 22 °C) melting of the T7 promoter becomes rate limiting, if at all.

A Revised Model for Promoter Recognition. The current results demonstrate quantitatively that the initiation of transcription by T7 RNA polymerase is relatively insensitive to the removal of large portions of the nontemplate strand downstream from position –6. Promoter specific binding, as reflected in the steady-state parameter K_m, is unchanged on removal of the nontemplate strand upstream to position –6, suggesting that the enzyme interacts with only the template strand in a form of the promoter which is melted in this region. In addition, the kinetic parameter k_{cat} appears to be dominated by a single process with an apparent activation energy of 26 kcal/mol, and promoter melting is not a major contributor to this kinetic barrier. Site-specific promoter recognition does not require a fully closed duplex promoter, as suggested by static footprinting data. Instead, we propose that the promoter is designed to allow energetically facile melting of this region of the DNA and that recognition of the promoter occurs in a structure in which the DNA is duplex from about position –6 or –7 upstream and is single stranded downstream of position –6. Both template and nontemplate bases contribute to recognition in the duplex domain, while only template interactions occur in the melted region.

ACKNOWLEDGMENT

We thank Charlie Schick for valuable comments and discussions and L. J. Maher for communicating results prior to publication.

SUPPLEMENTARY MATERIAL AVAILABLE

Steady-state kinetic fits for constructs listed in Figure 3, but not included in Figure 2 (1 page). Ordering information is given on any current masthead page.

REFERENCES

Basu, S., & Maitra, U. (1986) *J. Mol. Biol. 190*, 425–437.
Buc, H., & McClure, W. R. (1985) *Biochemistry 24*, 2712–2723.
Chamberlin, M., & Ryan, T. (1982) *Enzymes (3rd ed.) 15*, 87–108.
Chapman, K. A., & Burgess, R. R. (1987) *Nucleic Acids Res. 15*, 5413–5432.
Chapman, K. A., Gunderson, S. I., Anello, M., Wells, R. D., & Burgess, R. R. (1988) *Nucleic Acids Res. 16*, 4511–4524.
Gunderson, S. I., Chapman, K. A., & Burgess, R. R. (1987) *Biochemistry 26*, 1539–1546.
Ikeda, R. A., & Richardson, C. C. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3614–3618.
Maher, L. J., III (1992) *Biochemistry 31*, 7587–7594.
Martin, C. T., & Coleman, J. E. (1987) *Biochemistry 26*, 2690–2696.
Maslak, M., Jaworski, M. D., & Martin, C. T. (1993) *Biochemistry* (first paper of three in this issue).
McAllister, W. T., & Carter, A. D. (1980) *Nucleic Acids Res. 8*, 4821–4837.
McClure, W. R. (1985) *Annu. Rev. Biochem. 54*, 171–204.
Mecsas, J., Cowing, D. W., & Gross, C. A. (1991) *J. Mol. Biol. 220*, 585–597.
Metzger, W., Schickor, P., & Heumann, H. (1989) *EMBO J. 8*, 2745–2754.
Milligan, J. F., Groebe, D. R., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783–8798.
Muller, D. K., Martin, C. T., & Coleman, J. E. (1989) *Biochemistry 28*, 3306–3313.
Osterman, H. L., & Coleman, J. E. (1981) *Biochemistry 20*, 4884–4892.
Schick, C., & Martin, C. T. (1993) *Biochemistry* (second paper of three in this issue).
Schickor, P., Metzger, W., Werel, W., Lederer, H., & Heumann, H. (1990) *EMBO J. 9*, 2215–2220.
Strothkamp, R. E., Oakley, J. L., & Coleman, J. E. (1980) *Biochemistry 19*, 1074–1080.