Volume 8 Number 12 1980 Nucleic Acids Research

Alterations in two conserved regions of promoter sequence lead to altered rates of polymerase binding and levels of gene expression

James E.Stefano, Jill W.Ackerson and Jay D.Gralla

Biochemistry Division of the Chemistry Department and Molecular Biology Institute, University of California, Los Angeles, CA 90024, USA

Received 9 April 1980

ABSTRACT

Characterization of recombinant lac promoters highlights the importance of two regions of sequence conservation in promoters. The "Pribnow box" sequences are necessary for specific transcription in this system. This specificity is maintained when a mutated upstream sequence is introduced. However, changing the upstream DNA sequences influences both the rate of RNA polymerase binding in vitro and levels of expression in vivo.

INTRODUCTION

Two fundamental and interrelated questions concerning gene expression in E. coli are: What distinguishes promoter DNA from non-promoter DNA and what determines the maximal rate of promoter utilization? Since the answers lie ultimately in the sequence of DNA bases, much effort has been applied to determination of promoter sequences. One approach to extracting the relevant information is to compare promoters for conserved sequences which are likely candidates for functionally important promoter elements. This approach has led to the identification of two separated regions of moderate sequence conservation[1-6] centering near 10 base pairs ("downstream homology" or "Pribnow box") and 35 base pairs ("upstream homology") distal to the startpoint of RNA transcription (see Figure 1).

A second approach is to determine the location and effects of mutations which alter the function of an individual promoter. This approach has complemented the sequence comparisons since most mutations lie either in the upstream homology region or the downstream homology regions (see Ref. 6 for a review). In vitro studies of these mutant promoters have led to various proposals

for the role of sequences since the rate of promoter binding, level of transcription, or ability to initiate transcript may be affected by sequence changes within the two homology regions (Majors, unpublished work: ref. 7,8).

In order to elucidate the role of these regions *in vitro* and *in vivo* we have constructed, using recombinant DNA techniques, double mutants of the *lac* operon which contain a mutation in each region of sequence homology. These double mutants and their parent strains were studied both *in vitro* and *in vivo*. The *in vivo* study relied on a plasmid vector which we have constructed (Ackerson *et al.* manuscript in preparation) to allow expression *in vivo* of promoters created *in vitro*. Downstream homology sequences are found to be necessary for significant levels of specific transcription *in vitro*. However, a mutation upstream, when combined with either of two downstream sequences, leads to a reduced rate of promoter binding *in vitro* and a reduced level of gene expression *in vivo*. Thus, although the downstream sequences are the principal "mark" identifying these promoters, the upstream sequences also affect levels of gene expression, probably by contributing to the rate of promoter binding by RNA polymerase.

MATERIALS AND METHODS

Transcription Assay of Binding Rate: Lac promoter DNA was isolated as described previously[16] by endonuclease Eco Rl digestion of purified plasmid pMB9. *E. coli* RNA polymerase, approximately 50% active, was prepared and assayed according to the method of Burgess and Jendrisak[37] and catalyzes promoter-specific transcription as judged by criteria described previously[16]. The conditions for transcription were essentially as described in references 16 and 19, except that acetylated bovine serum albumin[35] (at 100 μg mL^{-1}) was used.

Binding rate experiments were initiated by mixing 2 picomoles of RNA polymerase and 1 picomole of purified promoter restriction fragment in a final volume of 25 μL at 37° C. At the times indicated heparin was added (to 100 μg mL^{-1}) to stop the binding reaction simultaneously with nucleoside triphosphates to

allow RNA production from productive complexes. ATP, GTP, and UTP were at 200 μM, and α-^{32}P-CTP at ca 15 Ci mmol^{-1} (prepared according to reference 38) was at 2 μM. RNA synthesis proceeded for 15 minutes at 37° C.

An equal volume of saturated urea containing marker dyes was added and an alliquot loaded directly on a 12% acrylamide gel containing 7M urea[34]. The lac-specific RNA was located on an autoradiograph, excised, and counted for Cerenkov radiation[19].

The construction of recombinant promoter strains and the cloning vehicle pAS21 is outlined briefly in the text and in detail elsewhere (in preparation).

Assay of Tetracycline Resistance: Tetracycline resistance was measured by a modification of the procedure of Tait et al.[23]. All measurements were done in E. coli strain MM 294. A freshly diluted transformed bacterial culture was grown to an A_{570} of 0.5 in LB broth. Triplicate aliquots were then diluted appropriately to give 100-200 colonies per plate. Each plate contained L agar (LB broth + 1.5% agar without added glucose) and a specific concentration of tetracycline. Plates were prepared immediately before use since tetracycline potency diminishes with time. After overnight incubation at 37° C colonies were counted and the three determinations averaged. Per cent survival was calculated relative to the number of colonies appearing at the lowest concentation of tetracycline.

RESULTS AND DISCUSSION

The Mutant Strains: The relevant lac promoter sequences are shown in Fig. 1. Class II promoter mutations are defined as those which exhibit decreased levels of expression in vivo both in the presence (crp$^+$, cya$^+$) and the absence (crp$^-$, cya$^-$) of functional catabolite activator protein[9]. These mutations have been sequenced[10] and are base changes within the consensus upstream homology region (see L305 in Fig. 1). Class III promoter mutations allow high level expression in the absence of functional catabolite activator protein[11,12]; two of these (see UV5 and p^S in Fig. 1) are changes which increase homology with the consensus downstream homology region (J. Gralla, J. Majors,

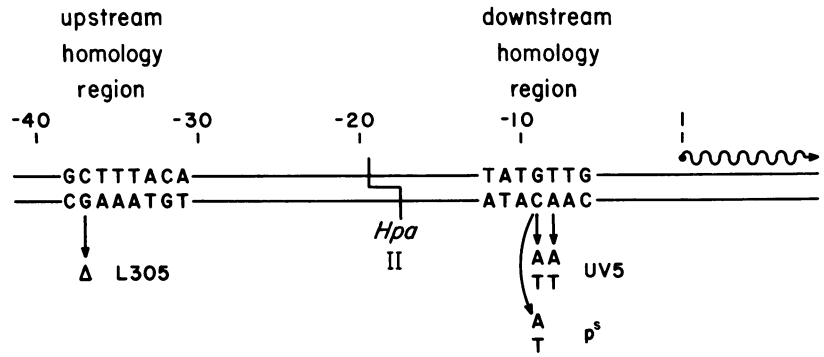

Figure 1. Schematic diagram of the lac promoters used. Distances marked are base pairs from the favored startpoint of the messages in vitro (shown by a wavy line). L305 is a "down" promoter mutant; UV5 and p^s are "up" promoter mutants (see text). The UV5 strain in addition contains the L8 CAP site mutation at position -66, not shown here. The Hae III restriction sites used to isolate the promoter fragments are located at positons +65 and -142. Conserved promoter sequences have been pointed out by several authors[1-6]. The wild-type sequence is from reference 15; the L305 sequence is from reference 10; see reference 3 for the UV5 and p^s changes.

quoted in Ref. 3).

We have taken advantage of the restriction enzyme cleavage sites within and about the lac promoter to construct recombinant promoters containing both class II and class III mutations. Specifically, endonuclease Hae III makes two cleavages which allow isolation of an intact promoter[13] while endonuclease HpaII cleaves within the promoter between the upstream and downstream homology regions. Thus, combined cleavage by these two enzymes allows isolation of an upstream half-promoter fragment of 120 base pairs and a downstream half-promoter fragment of 80 base pairs. Since individual mutant strains were available either on transducing phages[14] or plasmid vectors (J. Majors, F. Fuller, unpublished), an 80 base pair fragment containing either the UV5 or p^s sequence and a 120 base pair fragment containing the L305 sequence were easily isolated.

Briefly, the 120 base pair L305 fragment was joined to the 80 base pair UV5 or p^s fragment by T4 DNA ligase in order to obtain species which include the 200 base pair recombinant

promoter. These products were then ligated into the Eco Rl site of the plasmid vector pMB9[15]. Cells were transformed, plated on XG-tetracycline, and plasmids from blue colonies (an indication of cloned lac operator) were screened by restriction cleavage. A large fraction of these yielded 200 base pair inserts upon Eco Rl cleavage which when digested subsequently with Hpa II revealed the original 120 base pair fragments.

At this point we had available five mutant promoter strains for study; L305, UV5, p^S, L305 x UV5, L305 x p^S (see Fig. 1). Each strain is cloned and the promoters contained in otherwise identical restriction fragments. Such fragments can be used in a highly purified in vitro transcription system to assay for lac m-RNA initiation (19,16,7).

Figure 2 shows that the UV5 and p^S promoters support specific transcription in vitro, as evidenced by the appearance of the runoff lac RNA band on the polyacrylamide gel. By contrast, the L305 DNA fails to support efficient transcription of lac RNA. Thus, the behavior of these strains in vitro mimics their behavior in vivo (see below).

When either UV5 or p^S was crossed with L305, the recombinants supported specific transcription (Fig. 1) as evidenced by the preferential appearance of the same length lac RNA band produced by UV5 and p^S. Thus, the downstream sequences (UV5 or p^S) are dominant to L305 by this limited criterion; that is, their presence allows specific transcription even in the presence of a mutated upstream homology region. Since L305 alone failed to support specific transcription, the downstream sequences in the UV5 or p^S strains are also necessary. The experiment shown allows only a single round of transcription from productive compexes (see Legend). Since the ultimate level of transcription attainable (Table I) is nearly identical for the four competent promoters, recombination with the L305 mutation has at most a minimal effect on the ability of RNA polymerase to ultimately form a stable complex and initiate at promoters containing the UV5 or p^S sequences.

Rate of Promoter Binding: Although the class II promoter mutation L305 failed to alter the ability of RNA polymerase to bind to the p^S or UV5 promoters, we wished to see if it might

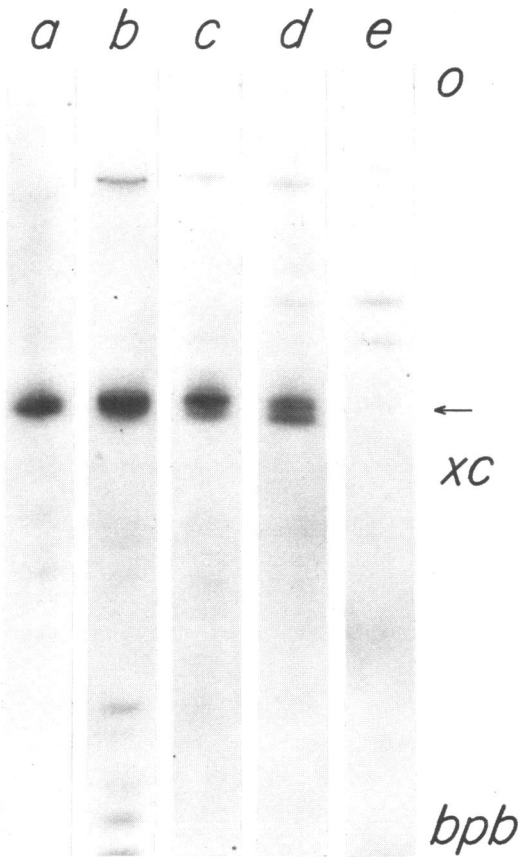

Figure 2. Denaturing gel analysis of mRNA synthesized in vitro from various lac promoter fragments. The arrow marks the lac specific transcripts as described previously[16,19]. a, UV5 (containing L8, see legend to Fig. 1); b, L305 x UV5; c p^s; d L305 x p^s; e, L305. O, origin; xc,bpb, marker dyes.

exert an effect instead on the rate of formation of productive complexes. Both class III lac mutant promoters[7] and phage fd promoters[18] exhibit variations in this rate of stable binding. It has been suggested that this kinetic variation may be more important than thermodynamic stability in determining the relative utilization of various promoters[18].

We used a transcriptional assay which is specific for promoter-bound transcriptionally competent complexes[19]. RNA

TABLE I

MAXIMAL TRANSCRIPTION LEVELS FROM MUTANT lac PROMOTERS in vitro

Strain	Yield (transcripts/promoter)
UV5	0.52
L305 x UV5	0.61
p^s	0.49
L305 x p^s	0.42

Conditions were as in Figure 2 except that the polymerase concentration was 200 nM. Titration data (unpublished) showed this concentration to be saturating for all four promoters under these transcription conditions. Specific activity of the CTP was 15.8 Ci $mmol^{-1}$; the Cerenkov counting efficiency was 40%.

polymerase was allowed to bind each fragment for various periods of time prior to addition of nucleoside triphosphates and heparin. Heparin inactivates any free RNA polymerase, thereby restricting bound, heparin-resistant RNA polymerase to a single round of RNA synthesis. Thus, the amount of the specific length lac RNA appearing on the polyacrylamide gel is proportional to the number of competent promoter-specific complexes formed at that time. An example of this assay is shown in Fig. 3.

Figure 4 displays the accumulation of functional complexes at the four promoters, UV5, L305 x UV5, p^s, and L305 x p^s. The introduction of the L305 mutation is seen to have a significant and consistent effect on the rate of productive complex formation at the lac promoters. The effect of L305 appears multiplicative in nature; that is, recombination of the L305 sequence with either of the downstream sequences, UV5 or p^s, results in approximately a three-fold reduction in binding rate as measured from the half-time for saturation. Thus, we can infer that upstream contacts altered by the L305 mutation are important determinants of the rate of promoter binding even though they are not important in determining the ability to ultimately form a bound complex.

The downstream contacts altered by UV5 must be important determinants of binding rate as well, as shown by the faster rate

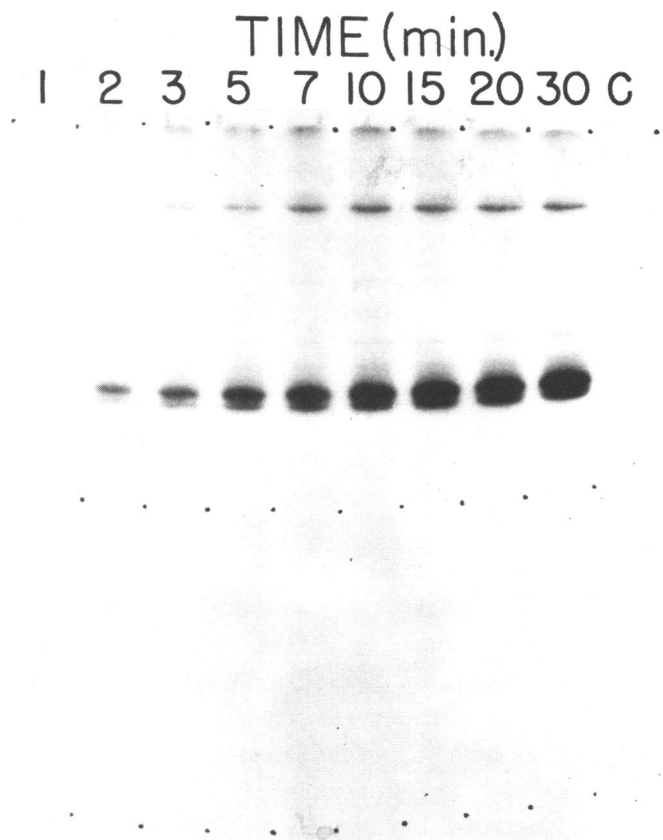

Figure 3. Binding rate experiment. RNA polymerase was allowed to bind the p^S promoter fragment for various periods of time shown prior to addition of heparin (to stop the binding reaction) and nucleoside triphosphates (which allows RNA synthesis from bound complexes). The dots mark positions of the origin, xc, and bpb. Lane "c" is a control experiment where DNA and heparin were added simultaneously to the polymerase.

of UV5 compared to p^S (a similar result has been obtained by J. Majors, unpublished). The UV5 mutation has already been shown to be required to attain tight binding and specific transcription, since wild-type sequences fail to either bind or transcribe efficiently *in vitro* in the absence of catabolite activator protein[19,20]. Taken together these data show that the base pair change represented by p^S is an important element in

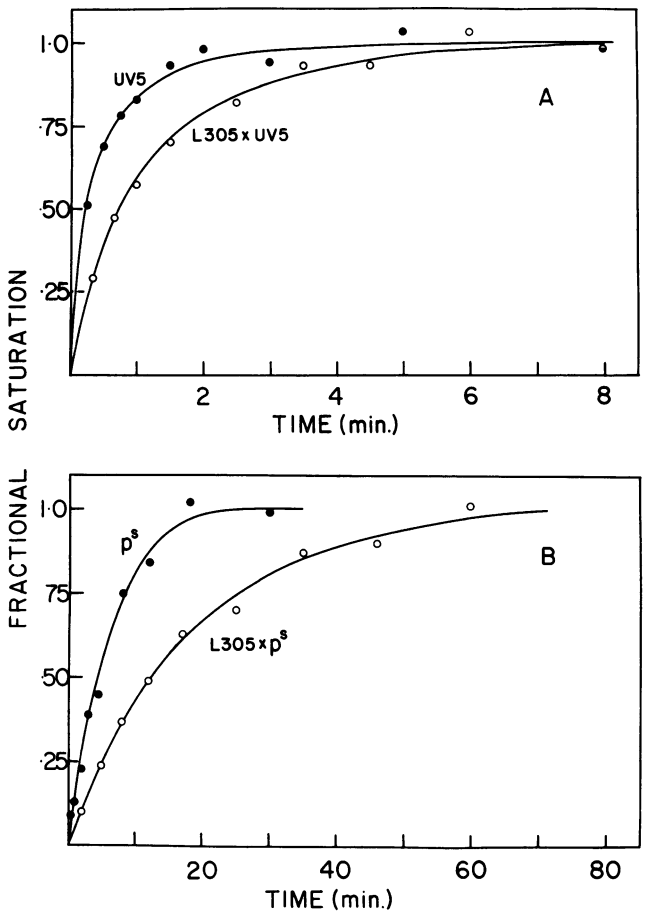

Figure 4. Kinetics of formation of transcriptionally competent complexes. Data from a typical experiment. Maximal binding has been normalized for each promoter. A.●, UV5; ○, L305 x UV5; B.●, p^s; ○, L305 x p^s. The yields of transcript in this experiment (RNA polymerase at 100 nM, lac promoter at 50 nM) relative to UV5 were L305 x UV5, 1.1; p^s, 0.43; L305 x p^s, 0.22.

transcriptional specificity and the second base changed in UV5 is important in accelerating the rate of promoter binding. Thus, the downstream homology region influences both the specificity and the rate of binding.

Promoter Directed Expression in vivo: We have shown that recombination of the upstream mutant sequence (L305) with either

of two competent downstream sequences (UV5 or p^S), results in a three-fold slowing of the rate of promoter binding *in vitro*. The physiological significance of these observations may be assessed by measuring levels of promoter expression *in vivo*.

We have constructed a vehicle which allows measurement of gene expression *in vivo* directed from promoters obtainable *in vitro* (Ackerson *et al.*, in preparation). We describe this system briefly. The plasmid pBR322[21] contains a single site for the restriction endonuclease Hind III which lies within the promoter for tetracycline resistance[8]. Substitution of an exogenous promoter sequence in this site results in a dependence upon this promoter for *tet* expression[22,39]. We have deleted upstream sequences in the vicinity of the Hind III site to eliminate *tet* promoter activity. Cells transformed with this vehicle fail to show any significant levels of tetracycline resistance.

Insertion of *lac* promoter fragments allows expression of the tetracycline resistance gene and restores promoter-dependent levels of tetracycline resistance to plasmid-containing cells. We have inserted the four competent promoters into this plasmid and measured the levels of lac promoter-dependent gene expression by assaying for tetracycline resistance *in vivo* by a modification of the method of Tait *et al.*[23].

Figure 5 displays the levels of tetracycline resistance conferred upon cells by the four promoters. Qualitatively, it is clear that insertion of L305 into either p^S or UV5 reduces the level of gene expression. As was the case for the promoter binding rates, the introduction of the L305 mutation leads to a significant and consistent inhibition. Therefore, even though L305 does not alter the ability to eventually form specific complexes with UV5 or p^S *in vitro*, it does interfere with promotion of gene expression *in vivo*. Quantitatively, the expression levels of UV5 and p^S are reduced comparably by recombination with L305. This multiplicative effect is analogous to the observed effect on the binding rates detailed above. The simplest interpretation is that the property of the L305 mutation which inhibits the rate of binding *in vitro* is the same property which leads to an inhibition of levels of gene expression *in vivo*.

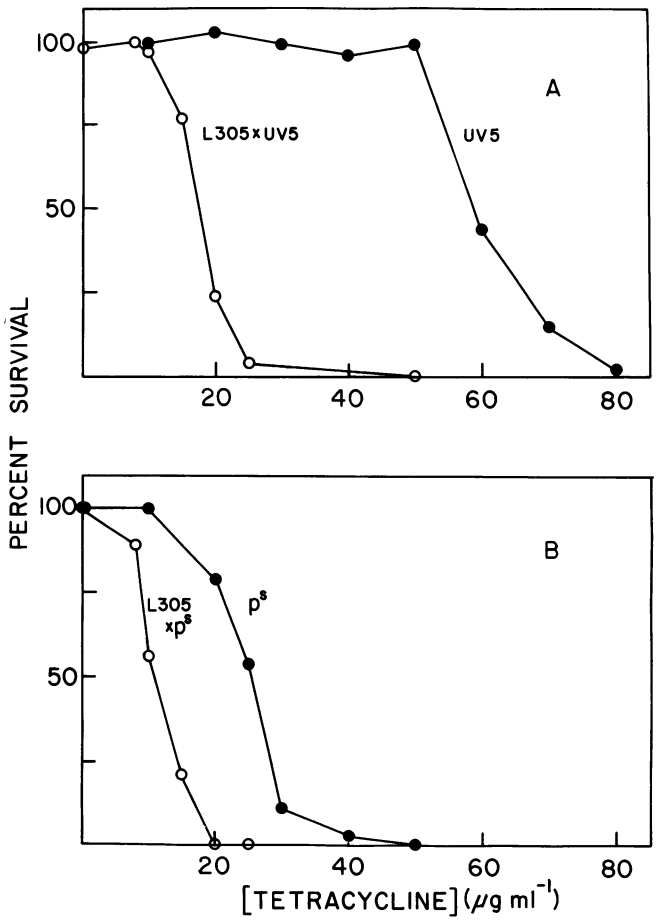

Figure 5. Lac promoter-dependent levels of tetracycline resistance. Measurements were done on E. coli strain MM294 containing a plasmid with a lac promoter substituted for the missing tetracycline promoter. The host plasmid pA21 exhibits 50% survival of 2 μg/mls of tetracycline (Ackerson et al., in preparation). The various lac mutant promoters UV5, L305 x UV5, p^s, and L305 x p^s lead to restoration of levels of tetracycline resistance as shown. Tetracycline resistance was measured by a modification of the procedure of Tait et al[23].

This conclusion is supported by the observation that the binding rate and levels of gene expression are reduced by a similar factor; both are reduced approximatly three-fold. This support is weakened somewhat as a consequence of the fact that

our in vivo conditions do not eliminate the stimulatory effect of catabolite activator protein. The p^S strain is known to be stimulated by this protein[24] and this may explain in part why the pair of p^S containing promoters are expressed in vivo at levels several-fold higher than would be predicted based on their binding rates relative to the pair of UV5-containing promoters. We expect this question to be resolved by studies in progress utilizing host strains deficient in functional catabolite activator protein. Nevertheless, the results showing that L305 leads to approximately three-fold reductions in the rate of p^S binding, the rate of UV5 binding, the level of p^S expression, and the level of UV5 expression, indicate that the rate of binding influences the level of expression, and that the sequences changed by L305 contribute to both processes.

It has been reported that for class III lac promoter mutants a correlation exists between the β-galactosidase levels and both the rate of promoter binding and the saturation level of complex formation in vitro[7]. All four competent promoters studied here ultimately achieve identical levels of complex formation with saturating polymerase (Table I), but vary widey in their rate of binding. Thus it is not the relative number of complexes that can be formed at promoters, but rather the differential rates at which promoter complexes form which consistently correlates with the levels of gene expression.

Promoter Utilization in vitro and in vivo: These experiments clarify the role of the downstream ("Pribnow box") and upstream homology regions in promoter function. Changes such as p^S and UV5 which increase downstream homology allow specific transcription and also contribute to increased rate of binding in vitro and increases in gene expression in vivo. The L305 mutation which occurs within the upstream homology region leads to a decreased rate of binding in vitro and a decrease in gene expression in vivo. These data confirm the functional importance of both regions and suggest that a crucial role of both is to determine the rate at which RNA polymerase may bind a promoter.

Initially, the upstream sequences were thought to be essential for the initial recognition of the promoter, in a step involving either melting of this region[25] or formation of a

complex utilizing these contacts exclusively[1]. Subsequently, several lines of evidence have reduced the necessity for these models involving entry sites. This paper and others[26,7,8] principally have demonstrated that alteration of sequences in this region does not preclude formation of the final pre-initiation complex. However, it has been suggested that this may be true only for a certain class of promoters[27]. The possibility that the recombinants shown here are members of such a class is made unlikely by the fact that the UV5 promoter is at least an order of magnitude faster in rate than the p^S promoter, yet does not show any significant relief from the dependence on these sequences in the binding process. At the other extreme, Rodriguez et al.[18] studying the tetraycline promoter, have raised the possibility that the upstream region plans no important role in promoter binding, but acts subsequently. We consider this possibility to be unlikely since the upstream-induced differences in binding kinetics we observe are sufficient to account for the in vivo behavior of the mutant lac promoters.

The rates of promoter binding (formation of productive complexes) we observe vary fifty-fold and are orders of magnitude slower than predicted[28,29] for diffusion-controlled reactions at these concentrations. Historically, this latter observation has been explained by one of two hypothetical mechanisms: (1) limited availability of free RNA polymerase due to its interaction with many non-promoter sites[29] or (2) formation of a heparin-sensitive intermediate complex at the promoter which must be slowly converted to the final bound complex[30]. The first hypothesis cannot alone account for the binding kinetics we observe since a fifty-fold variation in rate occurs among promoters contained within identically-sized restriction fragments. However, in a detailed kinetic study (Stefano and Gralla, in preparation) we have observed that the rate of productive complex formation depends directly upon the polymerase concentration but is not significantly altered by the promoter concentration for all four promoters presented here, a prediction made by the first hypothesis. Taken together, these data suggest that both mechanisms play a role in site selection at these promoters. However, the variation in the rate can only be accommodated by

the properties of the intermediate in mechanism 2 above. In this formalism, the inhibition effected by the L305 mutation is then due to a lower stability of this intermediate or its slower conversion to the final bound complex, or both.

Static studies of bound promoter complexes suggest that both the upstream and downstream homology regions are contacted by RNA polymerase[31,32,33,36]. Our results show that these regions set the rate of productive complex formation of lac promoters which correlates with gene expression in vivo. Further, we postulate that this rate is dependent on the properties of a kinetic intermediate at the promoter. The precise role of promoter DNA sequence in determining the attainable levels of promoter expression may only be revealed by describing the properties of this intermediate.

ACKNOWLEDGMENT

This research was supported by a USPHS traineeship to J.E.S. and research grants from the National Cancer Institute (Ca 19941) and the UCLA Academic Senate.

REFERENCES

1. Pribnow, D., J. Mol. Biol. 99, 419-443 (1975)
2. Schaller, H., Gray, C. and Herrmann, K., Proc. Natl. Acad. Sci. USA 72, 737-741 (1975).
3. Gilbert, W. in RNA Polymerase (eds. Losick, R. and Chamberlin, M.) 193-205 (Cold Spring Harbor Laboratory (1976)
4. Scherer, G., Walkinshaw, D. and Arnott, S., Nucleic Acids Res. 5, 3759-3773 (1978).
5. Siebenlist, U., Nucleic Acids Res. 6, 1895-1907 (1979).
6. Court, D. and Rosenberg, M., Ann. Rev. Genetics 13, 319-353 (1979).
7. Maquat, L. and Reznikoff, W., J. Mol. Biol. 125, 467-490 (1978).
8. Rodriguez, R., West, R., Heyneker, H., Bolivar, F. and Boyer, H., Nucleic Acids Res. 6, 3267-3287 (1979).
9. Hopkins, J., J. Mol. Biol. 87, 715-724 (1974).
10. Dickson, R., Abelson, J., Johnson, P., Reznikoff, W. and Barnes, W., J. Mol. Biol. 111, 65-75 (1977).
11. Silverstone, A., Arditti, R. and Magasanik, B., Proc. Natl. Acad. Sci. USA 66, 773-779 (1970).
12. Beckwith, J., Grodzicker, T. and Arditti, R., J. Mol. Biol. 69, 155-160 (1972).
13. Gilbert, W., Gralla, J., Majors, J and Maxam, A., in Protein-Ligand Interactions (eds. Blauer, G. and Sund, H.) 193-210 (Walter de Gruyter, 1975).

14. Reznikoff, W., Winter, R. and Hurley, C., Proc. Natl. Acad. Sci. USA 71, 2314-2318 (1974).
15. Rodriguez, R., Bolivar, R., Goodman, H., Boyer, H. and Betlach, M. in Molecular Mechanisms in the Control of Gene Expression (eds Nierlich, D., Rutter, W., Fox, C.) 471-478 (Academic, 1976).
16. Stefano, J. and Gralla, J., Biochemistry 18, 1063-1067 (1979).
17. Maizels, N., Proc. Natl. Acad. Sci. USA 70, 3585-3589 (1973).
18. Seeburg, P., Nusslein, C. and Schaller, H., Eur. J. Biochem. 74, 107-113 (1977).
19. Majors, J., Proc. Natl. Acad. Sci. USA 72, 4394-4398 (1975).
20. Reznikoff, W. in RNA Polymerase (eds. Losick, R. and Chamberlin, M.) 441-454 (Cold Spring Harbor Laboratory, 1976).
21. Bolivar, R., Rodriguez, R., Greene, P., Betlach, M., Heyneker, H. and Boyer, H., Gene 2, 75-93 (1977).
22. Neve, R., West, R.L. and Rodriguez, R. L., Nature 277, 324-325 (1979).
23. Tait, R. C., Rodriguez, R. L. and Boyer, H. W., Molec. Gen. Genet. 151, 327-331 (1977).
24. deCrombrugghe, B., Chen, B., Gottesman, M., Pastan, I., Varmus, H. E., Emmer, M. and Perlman, R. L., Nature 230, 37-40 (1971).
25. Dickson, R., Abelson, J., Barnes, W. and Reznikoff, W., Science 187, 27-35 (1975).
26. Okamoto, T., Sugimoto, K., Sugisaki, H. and Takanami, J. L., Nucleic Acids Res. 7, 2213-2222 (1977).
27. Reznikoff, W. S. and Abelson, J. N. in the Operon (eds. Miller, J. H.and Reznikoff, W. S.) 221-243 (Cold Spring Harbor Laboratory, 1978).
28. Alberty, R. A. and Hammes, G. G., J. Phys. Chem. 63, 274-279 (1958).
29. Hinkle, D. C. and Chambelrin, M. J., J. Mol. Biol. 70, 187-195 (1972).
30. Chamberlin, M. J., Ann. Rev. Biochem. 43, 721-775 (1974).
31. Simpson, R. B., Proc. Natl. Acad. Sci. USA 76, 3233-3237 (1979).
32. Johnsrud, L., Proc. Natl. Acad. Sci. USA 76, 5314-5318 (1978).
33. Schmitz, A. and Galas, D. J., Nucleic Acids Res. 6, 111-135 (1979).
34. Maniatis, T., Jeffrey, A. and van de Sande, H., Biochemistry 14, 3787-3794 (1975).
35. Gonzales, N., Wiggs, J. and Chambelrin, M., Arch. Biochem. Biophys. 182, 402-406 (1977).
36. Siebenlist, U. and Gilbert, W., Proc. Natl. Acad. Sci. USA 77, 122-126 (1980).
37. Burgess, R. R. and Jendrisak, J. J., Biochemistry 14, 4634-4638 (1975).
38. Reeve, A. E. and Huang, R.C.C., Nucleic Acids Res. 6, 81-90 (1979).
39. Windera, G., Gautier, F., Lindemmaer, W., Collins, J., Molec. Gen. Genet. 163, 301-305 (1978).