Volume 11 Number 4 1983 Nucleic Acids Research

The coupled use of 'footprinting' and exonuclease III methodology for RNA polymerase binding and initiation. Application for the analysis of three tandem promoters at the control region of colicin E1

Peter T.Chan[21] and Jacob Lebowitz[22]

Department of Microbiology, The University of Alabama, Birmingham, AL 35294, USA

Received 8 November 1982; Revised and Accepted 21 January 1983

ABSTRACT

In order to determine the initiation site for three promoters P1, P2 and P3 (5' to 3') in close proximity in the colicin E1 control region we developed a new methodology that couples ternary complex formation and the analysis of the 3' border protected from exonuclease III digestion. The initiation of transcription could be detected by measuring the shift in the position of the 3' protected border when RNA polymerase moved from its binary complex position to its ternary complex position. The latter stops at a specific nucleotide because transcription is initiated with one or more NTPs missing. This approach, coupled with "footprinting", can also be used to decide whether the formation of an RNA polymerase binary or ternary complex at one site excludes or weakens binding at neighboring sites. The location of 3' protected borders reveals the formation of respective binary and ternary complexes at non-saturating RNA polymerase conditions, whereas at saturating conditions only the distal 3' boundary is seen and exonuclease cannot penetrate further. However, if "footprinting" reveals proximal 5' patterns this establishes that simultaneous binding has occurred on the same DNA fragment. The data showed that this was true for P1 and P3 which are only 8 nucleotides apart. P2 could only be detected at non-saturating conditions since it overlaps both P1 and P3. The evidence from the literature and this study establishes P1 as the true colicin E1 promoter with the possibility that supercoiling may eliminate any role for P2 and P3.

INTRODUCTION

Recently Ebina *et al.* (1) published their DNA sequence of the promoter region of colicin E1, the *cea* gene of the plasmid *Col*E1. To locate the promoter of this gene RNA was synthesized *in vitro* on the Sma I-Hae III (570 bp) fragment that contains the putative control region for colicin E1 expression. From the analysis of the 5'-terminal sequences of the RNAs synthesized, the promoter and operator regions of the colicin E1 gene were assigned. This analysis revealed two promoters in close proximity: one containing a "SOS" operator sequence, strongly suggesting repression by the *lexA* protein, and the other located in the coding region for colicin E1. Very recently, additional data from the same laboratory established that

lexA binds to the "SOS" operator region repressing RNA synthesis from this proximal promoter (2). Furthermore, in a paper (3) dealing with the full nucleotide sequence of the colicin El gene it was indicated as unpublished data that the only in vivo transcript for this gene stems from the proximal, lexA repressible promoter and therefore the distal promoter does not appear to function in vivo.

In the course of our in vitro investigation of the colicin El promoter region we have independently confirmed the existence of the two promoters described above by quite different experimental approaches. The location of the promoters was established using DNAase I protection technique of Schmitz and Galas (4) which is commonly refered to as "footprinting". In order to determine the initiation sites for potential promoters and examine other features regarding possible RNA polymerase-RNA polymerase interactions, we developed new methodology which couples DNA sequencing, "footprinting" and exonuclease III analysis of the 3' protected borders of RNA polymerase binary and ternary complexes. The determination of the 3' protected borders refers to the nucleotide position or positions, 3' to the 5' ^{32}P labeled end of the DNA fragment used in the analysis, that are made resistant to exonuclease III digestion by RNA polymerase binding. These 3' protected borders are determined precisely by utilizing denaturing gel electrophoresis to measure the molecular weight of the partially digested DNA fragment. This molecular weight corresponds to the nucleotide length from the 5' labeled end to the 3' nucleotide position where exonuclease III stopped due to the obstruction presented by RNA polymerase. We first determined the 3' protected border of RNA polymerase-DNA binary complexes. Promoter specific initiation (5) with the appropriate set of NTPs results in a numerical shift of the nucleotide position of the 3' border due to RNA polymerase movement, from the initiation site to the nucleotide residue that is missing in the transcript sequence. The shift in the nucleotide position of the 3' border coupled with the DNA sequence yields initiation sites for in vitro transcription. In addition, for promoters in very close proximity, exonuclease III and DNAase I protection can be utilized together to detect whether RNA polymerase-RNA polymerase interactions occur by analyzing whether promoter occupancy at one site excludes binding at the other site. Furthermore, we may ask whether upstream 5' ternary complex formation and subsequent RNA polymerase movement affect a downstream RNA polymerase binary complex.

The utilization of the colicin El promoter region is a good model

system to show the applicability of the experimental techniques outlined above. In order to do this we must present the DNA sequence of the colicin E1 promoter region as determined by us since a different restriction fragment was used for DNA sequencing by Ebina et al. (1). In terms of the DNA sequence there is an exact correspondence between the data. The protection data will be presented to emphasize our experimental rationale hence referral to Ebina et al. (1) will be minimal in the results section but expanded in the discussion section.

MATERIALS AND METHODS

Bacterial strains. Growth of E. coli K12, χ2180 (from the collection of R.C. Curtiss III) and ColE1 DNA purification were described previously (6).

Enzymes. All restriction endonucleases and exonuclease III were purchased from New England Biolabs. E. coli alkaline phosphatase and DNAase I were purchased from Worthington (Millipore Corp.). T4 polynucleotide kinase was purchased from P. L. Biochemicals. RNA polymerase was prepared as described previously (6).

DNA sequence analysis. Restriction maps of the Msp I d fragment for AluI, Fnu4H I, Hha I and Rsa I were obtained in order to select appropriate overlap fragments for the sequence analysis (Chan and Lebowitz, manuscript in preparation). Isolation and 5' end labeling of respective fragments were performed as previously described (6). DNA sequencing was carried out by a slight modification of the method of Maxam and Gilbert (7) as described previously (6).

Characterization of RNA polymerase-DNA interactions. RNA polymerase "footprinting" was carried out according to the method of Schmitz and Galas (4) using the appropriate 5' ^{32}P end labeled fragment that was used for the sequence determination. The amount of RNA polymerase needed to saturate the promoter sites was determined by evaluating the DNAase I protection pattern. This was accomplished by adding 2, 4, 8 and 16 µg of RNA polymerase by weight to labeled DNA (Hha I large subfragment) in the range of 1-10nM (DNA phosphate); after 10 min, pancreatic DNAase was added to five different aliquots to a final concentration of 1 µg/ml for incubation times of 0.5, 1.0, 1.5, 2.0, and 3.0 min. at 37°C. Reactions (final volume 20-25 µl) were stopped and pooled and treated as described (4); autoradiograms were inspected for the appearance of "footprints". Ternary complex assay was carried out according to the method of Taylor and Burgess (5).

Protection by RNA polymerase against E. coli exonuclease III digestion was carried out according to the method of Shalloway et al. (8). Exonuclease III was added to 5' ^{32}P end labeled DNA-RNA polymerase binary or ternary complexes, initiated with A+U or G+C, under RNA polymerase saturating (10 µg) or non-saturating, (2 µg) conditions using 50 units of exonuclease for the former and 25, 50 and 100 for the latter. Reactions were incubated at 37°C for 30 min.

RESULTS

Nucleotide sequence of the Msp I d fragment of ColE1 DNA. The nucleotide sequence of the Msp I d fragment is presented in Fig. 1. There exist within this fragment several potential promoter sequences. Therefore further experiments were necessary to identify the true cea promoter.

RNA Polymerase "Footprinting" and Ternary Complex Assays. The DNAase I protection patterns for RNA polymerase complex formation can be seen on one sequence gel if a subfragment of Msp I d is used. The Hha I-Msp I fragment (-167 to 212) was obtained by isolating the Hha I e fragment of ColE1 DNA (-167 to 277); this fragment was then 5' end labeled and recut with Msp I to generate the -167 to 212 subfragment used for the analysis. By varying the amount of RNA polymerase in the assay we can demonstrate that the promoter region of the cea gene contains two tandem RNA polymerase binding sites protecting the regions between nucleotide number -137 to -48 and from -40 to the +30 region respectively, Fig. 2. The latter protection occurs at lower polymerase concentrations indicating that the association constant is greater for the distal RNA polymerase-DNA complex relative to the proximal interaction.

In order to establish whether these RNA polymerase interactions are capable of initiating transcription we used different combinations of NTPs and tested for the formation of ternary complexes (four nucleotide RNA chain or greater) stable in 0.5M KCl that are retained on nitrocellulose filters (5). The ability to discriminate potential initiation sites is increased with sets of only two NTPs. We found that strong ternary complexes could be formed with G+C and A+U (data not shown). We also tested single NTPs and found that only U allowed the formation of a weak ternary complex (data not shown). An examination of the sequence (Fig. 1) shows that the ternary complex data are consistent with only four possible initiation sites. For the G+C case, initiation occurs at the +10 G position and synthesizes GCGG. A perfect consensus "Pribnow" box TATAATG (-4 to

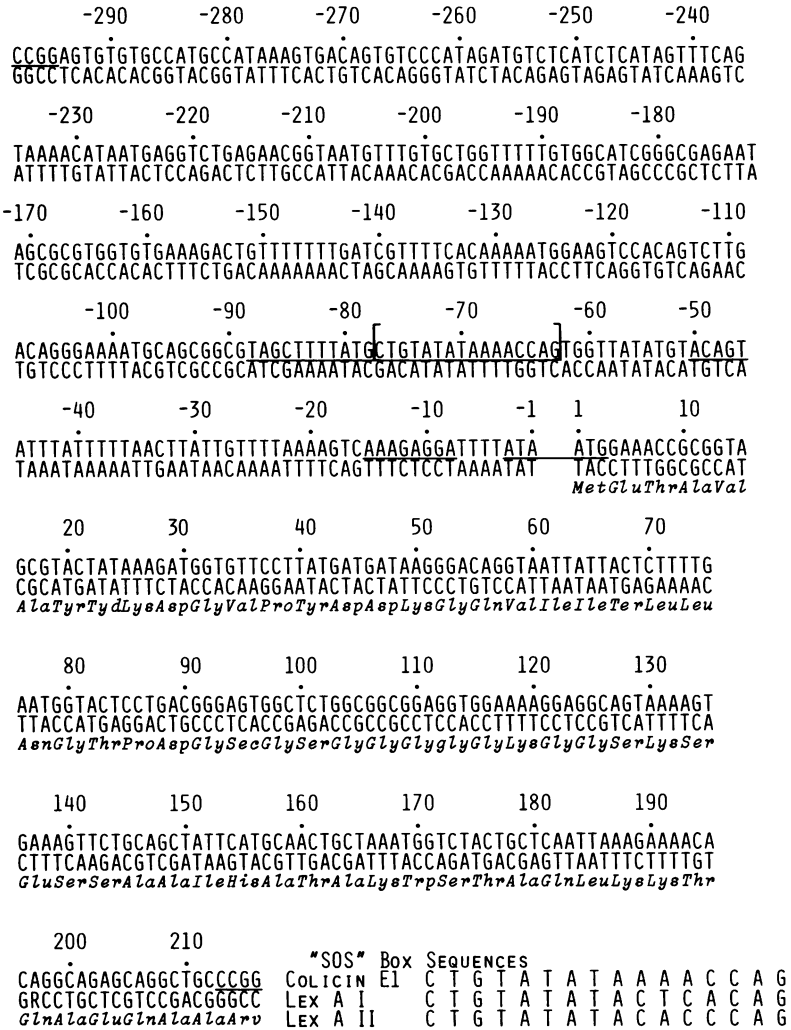

Figure 1. The sequence of the Msp I d fragment of ColE1 DNA. The top strand is the coding strand 5' to 3'. The sequence is numbered with respect to the amino terminal residue of colicin E1 with positive numbers 3' and negative numbers 5'. The underlined sequences from 5' to 3' are as follows: The 5' Msp I site; two overlap "Pribnow" boxes TAGCTTT and TTTTATG for the P1 promoter; "SOS" operator sequence bracketed; Pribnow box for P2; Shine-Dalgarno sequence; "Pribnow" box for the P3 promoter; the 3' Msp I site. The first twenty amino acids were determined by protein sequencing data kindly supplied to us by Sauer, R.T., Suit, J.L. and Luria, S.E. and by Cramer, W.A. independently. The remaining amino acids were determined from the nucleotide sequence. At the end of the Msp I d fragment we present several "SOS" operator sequences. This allows comparison of the sequence homology of the "SOS" box of colicin E1 with the two lexA "SOS" boxes.

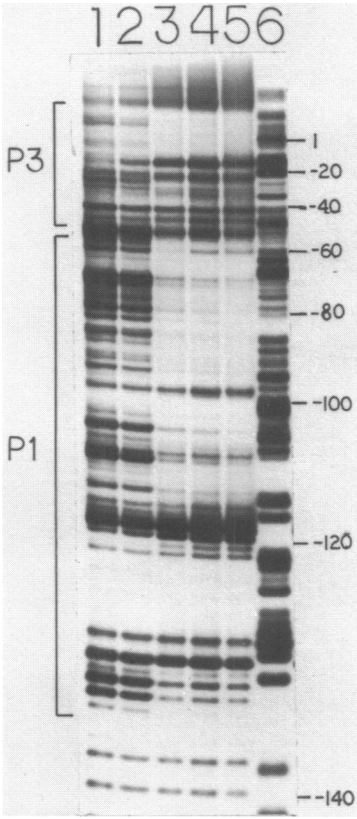

Figure 2. Binding of RNA polymerase to the Msp I subfragment of Hha I e (-170 to 216). 5' ^{32}P end labeled samples were treated with DNAase I (1 µg/ml) in the absence (lane 1) and the presence of RNA polymerase (lanes 2-5). The respective amount of RNA polymerase added was 2, 4, 8 and 16 µg (lanes 2-5) followed by the addition of DNAase I to detect protected or enhanced phosphate sites. Lane 6 contains the same DNA sample cleaved chemically at purine residues using the procedure that yields A > G. The figure is numbered in accordance with the numbering of the sequence (Fig. 1) and the protected promoter regions for P1 and P3 are indicated by large and small brackets respectively.

+3) occurs before the initiation site. This initiation data is completely in accord with the P3 footprint. For the A+U case, initiation occurs at the -73 A residue and synthesizes AUAUAAAA. Potential Pribnow boxes precede the -73 A initiation site. One possibility is TTTTATG (-84 to -78) which is almost identical to the "Pribnow" sequence of the tet promoter (9). Another possibility is the overlap sequence TAGCTTT (-88 to -82) which is similar to the λ pRM promoter (9). Of considerable significance in identifying the correct cea promoter is the appearance of an "SOS" box (10-15) CTGTATATAAAACCAG (-77 to -62) which would block initiation at position -73 if lexA binds at this operator sequence (Fig. 1). The -73 initiation site is in accord with the footprint designated P1.

Alternative possibilities for A+U initiation exist at positions -45 and -41. The former would generate the transcript AUUUAUUUUUAA while the latter would form AUUUUUAA. There are potential "Pribnow" boxes for both

initiation sites: TTATATG (-58 to -52) and TACAGTA (-51 to -45) for initiation at -45 and -41 respectively. The observation that U alone can initiate transcription (data not shown) suggests that a secondary start occurs at position -42 and generates the transcript UUUUU allowing the formation of a weak ternary complex that we can detect by filter binding. Usually a transcript needs to be extended four nucleotides to stabilize a ternary complex (5), consequently secondary initiation at position -40 is more likely than at position -44 since the latter would generate only a UUU transcript. However, we can not rigorously exclude initiation from position -44 as the source of a weak U generated ternary complex. The fact that U is capable of supporting limited transcription raises the possibility of another promoter between P1 and P3, which we tentatively designate P2. Clearly, additional evidence is required to establish the existence of P2.

P3 forms the strongest in vitro binding site on the Hha I large subfragment; however it cannot possibly be the true colicin E1 promoter since it initiates RNA synthesis in the cea structural gene. The P3 "Pribnow" box contains the AUG start codon for protein synthesis (Fig. 1). This information is firmly based upon the work of Ebina et al. (1) as well as the knowledge of the first twenty N-terminal amino acids of colicin E1 from two laboratories (Sauer, R.T., Suit, J.L. and Luria, S.E. and Cramer, W.A., personal communication) and rules out the distal promoter's role in colicin E1 synthesis. The protein sequence is in full agreement with the DNA sequence.

Consequently, the "footprinting" and ternary complex data support three in vitro tandem promoters, designated P1, P2 and P3 in the 5' to 3' direction on the coding strand with the highly unusual situation that the strongest binding site P3 covers part of the colicin E1 structural gene. Since P1 and possibly P2 appear to initiate with A+U we can not distinguish by the ternary complex assay whether both promoters are utilized since both would form salt resistant complexes and be retained on nitrocellulose filters. Consequently, the ternary complex assay cannot discriminate promoters on the same fragment that initiate with identical NTPs.

In order to resolve the initiation sites from different promoters in close proximity, we utilized exonuclease III digestion (8) to partially digest RNA polymerase-DNA fragment complexes to the point where exonucleolytic attack stops due to the obstructions generated by the molecular domain of RNA polymerase. If the 5' end of the coding strand of the duplex

fragment is ^{32}P labeled, we will detect the 3' barriers to exonuclease III digestion, generated by RNA polymerase, distal to the "Pribnow" box as exonuclease III attempts to digest this strand in a 3' to 5' direction. To detect initiation from a particular promoter we analyzed the 3' protected borders of binary and ternary complexes. The movement of RNA polymerase corresponds to the size of the short transcript formed in the ternary complex, and this change should be reflected in a new 3' protected border. Hence, initiation is detected as a shift in the position of the 3' protected border and this is measured as a difference in molecular weight of the exonuclease III digested binary and ternary complexes since unique molecular weight fragments of the labeled coding strand will be generated by exonuclease III digestion of the respective complexes. To extend this analysis to a DNA fragment containing multiple promoters, such as the colicin E1 control region, requires that we consider possible complications. For example, if the most distal 3' promoter is occupied with RNA polymerase this will block the access of exonuclease III to upstream promoters. In terms of the colicin E1 system, exonuclease III will stop at the RNA polymerase bound at P3, masking the existence of any RNA polymerase complexes at P1 and P2. Consequently, to detect potential 3' barriers to exonuclease III generated by P1 and P2 binary and ternary complexes we must modify the assay conditions to dissociate RNA polymerase from the P3 site enough to allow exonuclease III to reach upstream binary and ternary complexes.

We will show below that at high RNA polymerase concentrations initiation at P1 forces sufficient dissociation of RNA polymerase from P3 to detect the 3' protected boundary for the ternary complex of P1. A more useful approach is to lower the RNA polymerase concentration, generating sufficient one to one RNA polymerase-DNA complexes to allow detection of all existing 3' protected boundaries in one experiment. In the experiments described below we utilized "footprinting" and the exonuclease III assay in combination, allowing for a detailed analysis of RNA polymerase binding and initiation at P1, P2 and P3 under high and low RNA polymerase concentration.

Detection of promoter initiation sites by protection against exonuclease III digestion of binary and ternary complexes at RNA polymerase saturating conditions.

If RNA polymerase binding at P3 excludes binding at P1 or vice versa full DNAase I protection of either site would be impossible. Consequently,

the "footprinting" data indicates that RNA polymerase binding at P1 and P3 occurs simultaneously since full protection of both sites are generated at high RNA polymerase concentration. On the other hand, the location of P2 strongly suggests that binding cannot occur at this site if either P1 or P3 are occupied. To test for the detection of initiation from P1 and P3 we employed exonuclease III digestion of binary and ternary complexes coupled with a reexamination of the "footprint" patterns. To insure full simultaneous binding at P1 and P3 we used saturating amounts of RNA polymerase. This would exclude binding at P2.

The results of Fig. 3 confirm simultaneous binding at both promoters. This is seen by examining lane 7 which shows only one narrow set of three protection boundaries corresponding to the 3' end of P3. However, under identical conditions of RNA polymerase binding, "footprints" are generated at both P1 and P3 (lane 3). Consequently as predicted, RNA polymerase binding at P3 prevents the detection of the simultaneous binding of another RNA polymerase molecule at P1. If P1 were occupied and P3 free on some DNA fragments we would observe a 3' protection boundary for the P1 promoter. This is indeed seen as a faint band in lane 7. However this can not represent the total binding at the P1 site since we observe a strong P1 "footprint" (lane 3). Consequently most P1 binding is masked by the simultaneous binding at the P3 site.

Given the close spacing of eight basepairs between "footprint" boundaries we would envision that limited initiation at P1 could promote the dissociation of the enzyme occupying P3 and generating an accessible 3' region for exonuclease III until it reached the ternary complex of P1. If our initiation assignment is correct for P1 then transcription with A+U only should generate a movement from -73 to -66 (Fig. 1). The results of this experiment are presented in lane 8 of Fig. 3. We now clearly see an exonuclease III boundary at the 3' end of P1 which is displaced by 7 or 8 basepairs from the faint P1 binary boundary of lane 7. Consequently, RNA polymerase transcription at P1, generating an eight nucleotide ternary complex, promotes removal of the enzyme at P3. This allowed exonuclease III to digest through those fragments in which RNA polymerase dissociated from P3. An examination of the "footprint" data for A+U initiation (lane 4) reveals significant changes throughout the P1 region with changes in DNAase I susceptibility at the -40 region. It is obvious from the exonuclease III results in lane 8 that A+U initiation promotes only partial dissociation of P3 to unmask P1 since intense P3 protection boundaries are

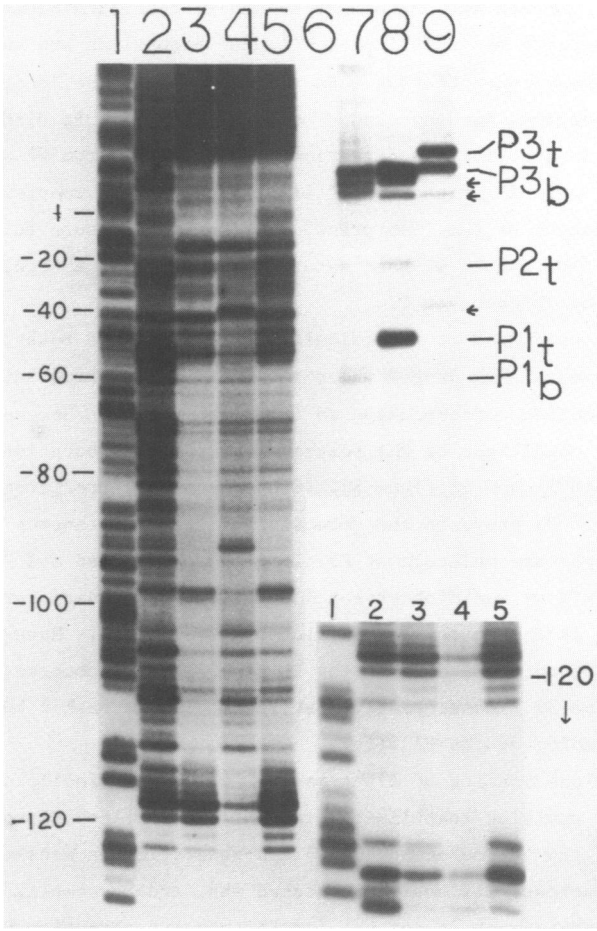

Figure 3. DNAase I and exonuclease III protection patterns ("footprints" and 3' borders) of binary and ternary complexes formed under RNA polymerase (10 µg) saturating conditions. 5'^{32}P end labeled Hha I fragment samples were treated as follows: Chemical (A>G) and DNAase I (1.0 µg/ml) cleavage in the absence of RNA polymerase respectively (lanes 1 and 2). DNAase I cleavage of the binary complex (lane 3) and the ternary complexes formed by initiation with A+U (lane 4) and G+C (lane 5). Protection patterns for residues 5' of -120 are shown in the lower right hand corner of the figure. Exonuclease III digestion (50 units) of free DNA (lane 6), the binary complexes (lane 7), the A+U ternary complexes (lane 8), and the G+C ternary complexes (lane 9). RNA polymerase-DNA exonuclease III protected complexes are designated with a subscript b or t to indicate a binary or ternary complex 3' protection boundary respectively for the P1, P2 and P3 promoters where appropriate. Arrows indicate internal secondary boundaries.

still present. In addition we also observed in lane 8 a faint band at the P2 promoter site. This would represent a P2 ternary complex since P2 initiation also occurs with A+U.

Finally, the presence of G+C initiation should only occur at P3 forming a ternary complex of four nucleotides +10 to +13 (Fig. 1). This should prevent exonuclease III accessibility to the P2 and P1 regions. As seen in lane 9 of Fig. 3, G+C initiation shifts the 3' protection boundary for P3 forward by 4-5 basepairs as predicted from the initiation assignment and sequence. Lane 5 shows the "footprint" for the G+C ternary complex and it can be seen from comparison with the binary complex "footprint" data (lane 3) that new DNAase I cleavages occur only in the P3 region (-10 to +10). However we still observe a strong binary "footprint" for P1. As expected the P3 ternary complex masks binding at P1.

Analysis of RNA polymerase protection under non-saturating conditions.

To resolve the 3' protected border of the P1 binary complex and confirm both the P1 initiation site and the new 3' border of the P1 ternary complex, we lowered the RNA polymerase concentration five fold in order to promote dissociation from P3 and allow exonuclease III to penetrate to the respective upstream protected sites. If RNA polymerase binding and initiation also occurs at P2, we should be able to detect a new 3' barrier to exonucleolytic digestion corresponding to protection at this site. Consequently, at non-saturating RNA polymerase conditions we should be able to see all 3' barriers to exonuclease III digestion. This is indeed the case as shown in Fig. 4. Lanes 6-8 represent the formation of A+U initiated ternary complexes and binary complexes (P3) subjected to 100, 50 and 25 units of exonuclease III respectively. In these lanes we observe the following: the binary complex for P3 with two minor bands due to further internal 3' exonucleolytic cleavage; the P2 ternary complex (lane 8) which shifts position by 1-2 nucleotides and diminishes as the exonuclease III concentration increases. As the first P2 ternary complex boundary diminishes a second internal boundary increases (lanes 6,7). Finally we observe the P1 ternary and binary complexes (lane 8) with the latter disappearing with increasing exonuclease III concentration (lanes 6,7) indicating a much weaker binary complex.

In regard to P2 and P3, it appears that exonuclease III is first blocked by RNA polymerase at a sharp barrier. If this blockage can be overcome the exonucleolytic attack proceeds rapidly to the next barrier generating additional bands. The two P1 bands correspond to the predicted separation of

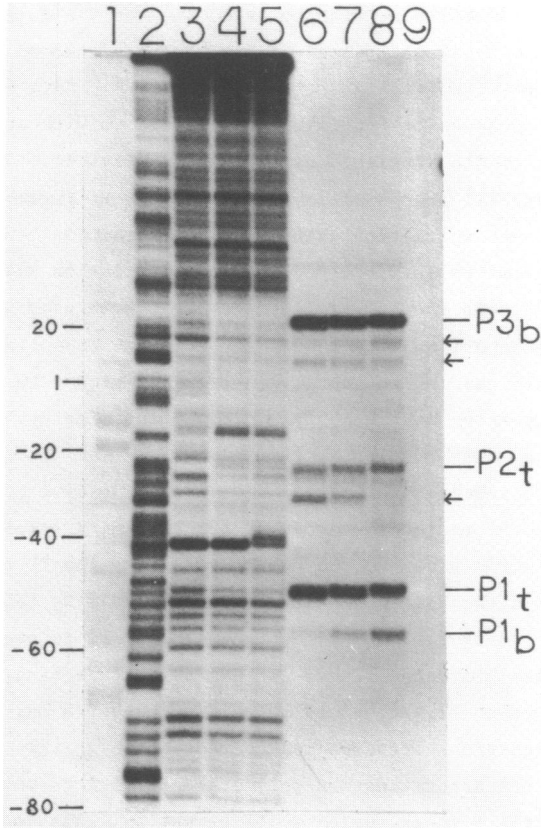

Figure 4. DNAase I and exonuclease III protection patterns ("footprints" and 3' border) of binary and ternary complexes formed under RNA polymerase (2 µg) non-saturating conditions. 5' ^{32}P end labeled Hha I fragment samples were treated as follows: Cleavage at purine residues (A>G), faint bands, pyrimidine residues (C>T), dark bands, and DNAase I (1.0 µg/ml) digestion in the absence of RNA polymerase respectively (lanes 1-3). DNAase I cleavage of the binary complexes (lane 4), and the ternary complexes initiated with A+U (lane 5). Exonuclease III digestion (50 units) of free DNA (lane 9) and the A+U initiated ternary complexes with 100, 50 and 25 units respectively (lanes 6-8). RNA polymerase-DNA exonuclease III protected complexes are designated with a subscript b or t to indicate a binary or ternary complex 3' protection boundary respectively for the P1, P2 and P3 promoters where appropriate. Arrows indicate internal secondary boundaries.

binary and ternary complexes of 8 nucleotides. The detection of a P1 binary boundary means that a fraction of the complexes failed to initiate transcription. Under these circumstances RNA polymerase dissociation from P1 would allow complete exonuclease III digestion at elevated exonuclease

levels causing the disappearance of the binary band. A partial failure to initiate transcription is also observed for P3 (Fig. 3, lane 9). The data strongly suggest that a stable binary complex does not form for P2, i.e. initiation must occur to stabilize the interaction at this site. Hence non-initiated RNA polymerase must rapidly dissociate from the P2 site.

It can be seen from Fig. 4 that the 3' boundaries determined by DNAase I protection for P3 and P1 are different than the 3' boundaries protected against exonuclease III digestion. Exonucleolytic attack digests 7 and 10 nucleotides more into the P3 and P1 promoter regions respectively than DNAase I, defining different 3' borders than "footprinting". In addition, significant differences in DNAase I sensitivity can be detected for P1 between binary and ternary complexes. Of note is a simultaneous increase and decrease of DNAase I sensitivity at the nucleotide residues at the -40 region. We believe that these changes cannot be accounted for by those fragments binding RNA polymerase and initiating at P2 since binding and initiation at P1 is much greater. In addition Fig. 3 (lane 4) shows a change at the -40 position for A+U initiation under RNA polymerase saturating conditions where binding at P2 appears virtually excluded. Consequently, the changes observed at the -40 region must be due to ternary complex formation at P1. In Fig. 2 we estimated that the P1 3' boundary ended at position -48; consequently, movement of RNA polymerase due to A+U initiation would shift the 3' boundary to position -40, supporting the interpretation that the P1 ternary complex produced changes in DNAase I sensitivity in that region. It is apparent from Fig. 4 that exonuclease III can digest further than the -40 position to approximately position -50. Consequently, for P1, ten nucleotides are susceptible to exonucleolytic attack that are resistant to DNAase I. For P3, seven nucleotides are susceptible. These differences in nuclease susceptibility probably reflect the way DNA strands interact with regions of RNA polymerase and further study could yield important structural features.

In conclusion, we observe all barriers to exonuclease III digestion for P1, P2 and P3 RNA polymerase-DNA complexes with the exception of the P2 binary complex.

DISCUSSION

The utilization of exonuclease III 3' border protection coupled with "footprinting" offers a powerful probe of promoters in close proximity. It also provides an excellent method for detecting initiation sites using ter-

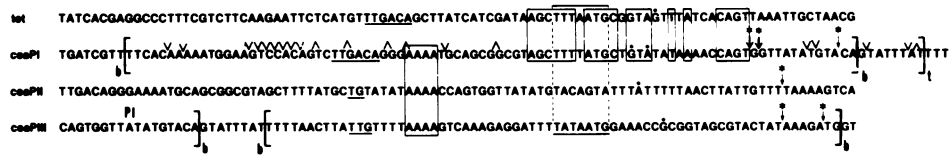

Figure 5. The nucleotide sequences surrounding the three in vitro promoters found in the control region of colicin E1. These sequences are aligned at their "Pribnow" boxes (vertical dashed lines) as well as with the tet promoter of pBR322 for comparison purposes (see text). The horizontal bracket represents the sequence homologous to known "SOS" boxes for lexA binding. The vertical brackets delineate the RNA polymerase binding sites ("footprints") based upon DNAase I protection with the b and t subscripts indicating binary or ternary complex formation respectively. Differences in DNAase I sensitivity between binary and ternary (A+U) complexes are indicated: Λ, sensitivity increased; V, sensitivity decreased. Initiation sites are shown with a filled dot above the site. Asterisks with downward arrows indicate barriers to exonuclease III digestion for binary, and ternary complexes, 5' and 3' respectively. P2 shows only the exonuclease III boundary for the ternary complex. The -10 and -35 regions of the promoters are underlined and the boxed sequences indicate homology between promoters.

nary complex formation.

In Fig. 5 we present the sequences of P1, P2 and P3 and the tetracycline (tet) promoter for comparative purposes. The brackets for P1 and P3 represent the protection borders from the "footprint" analysis. We do not have definitive protection boundaries for P2 since "footprint" analysis was always performed with strong P1 and P3 backgrounds. Although we have indicated two possible P2 "Pribnow" boxes, and initiation sites spaced three nucleotides apart, we favor the -41 site since it is most consistent with the formation of a ternary complex initiating with U only. This is indicated in Fig. 5 without reference to the -45 site, although we can not exclude the latter without further experimental data. The "Pribnow" box for P2 in Fig. 5 is part of the protected sequence of P1 at the 3' binary border, and the P2 initiation site is two nucleotides from the 5' binary border of P3. Clearly RNA polymerase cannot bind at P2 unless both P1 and P3 are free. The exonuclease III data provided strong evidence that this is indeed the case under RNA polymerase non-saturating conditions. In contrast we have shown that P1 and P3 can bind simultaneously to the Hha I-Msp I restriction fragment. The two promoters are separated by eight nucleotides and consequently A+U initiation at P1 would move the P1 RNA polymerase in direct contact with the P3 RNA polymerase molecule. This promotes RNA polymerase dissociation from P3 or we would not have detected the P1 ternary complex 3' border. This is accompanied by both increased

and decreased DNAase susceptibility (Figs. 3-5). Of particular interest is the appearance of a new cleavage site at -40 which is the 3' border of the P1 ternary complex.

We again note that exonuclease III can digest ten and seven nucleotides for P1 and P3 respectively that are resistant to DNAase I attack. In light of the differences between the protection boundaries of an endo and exonuclease, it would be of interest to examine whether differences exist in protected and enhanced phosphodiester sites, using other methods for cleaving DNA beside DNAase I. The variations in protection patterns may allow us to resolve how accessible or inaccessible various sites are to cleavage.

It is, of course, of interest to compare our results by the "footprinting" and new exonuclease III methodology with the results of Ebina et al. (1) in which in vitro transcripts were sequenced to determine initiation sites. In regard to in vitro promoters, both approaches identified P1 and P3. However, the following differences in the results should be noted: Ebina et al. (1) found that initiation occurred at the -75 G residue instead of the -73 A residue and they concluded that the "Pribnow" box for P1 is TAGCTTT instead of the overlap sequence TTTTATG. We accept the -75 G initiation site and their P1 "Pribnow" box assignment for the following reasons: previous data showed that only one transcript is synthesized in vitro using γ^{32}P ATP (16) and this initiation site is located at 0.12 map units in the copy number control region (6). Consequently P1 does not normally initiate with an A residue. We view this difference of two residues in initiation as quite minor since many promoters can initiate using different neighboring purine nucleotides (9). When transcription is initiated with a full complement of NTPs, the -75 G residue overwhelmingly predominates, whereas, if only A+U is used to initiate transcription, RNA polymerase selects the -73 A as a strong secondary initiating nucleotide. We undoubtedly selected a strong secondary start for P1 using A+U since reexamination of A, G, and U initiation revealed a stronger ternary complex than the A+U complex [data not shown]. Under the latter conditions, RNA polymerase would still stop at residue -65 due to the missing C needed to continue transcription. Consequently, we would generate an essentially identical ternary complex if we had initiated with A+G+U, with the RNA two nucleotides longer. The ability of RNA polymerase to initiate at different sites for the same promoter does point out that one must be cautious in assigning the initiation site from the ternary

complex-exonuclease III methodology if two or more potential A and G residues are capable of initiating transcription. Consequently, one should test the relative strength of initiation from the intensity of the protected 3' boundary using the appropriate sets of NTPs as a possible method for discrimination of different initiation sites from the same promoter. This, coupled with the strength of ternary complex formation as developed by Taylor and Burgess (5), should differentiate between primary and secondary initiation sites. When we review our data on this matter for P1, we conclude again that primary initiation occurs at the -75 G residue as found when a transcript is sequenced from that promoter (1). This conclusion makes the TAGCTTT more favorable as the P1 "Pribnow" box since TTTTATG would be too close to the initiation site.

P2 and the simultaneous binding of RNA polymerase at P1 and P3 were not detected by Ebina et al. (1). In their analysis, however, no attempt was made to study RNA polymerase-DNA interactions. Consequently our study expands on the in vitro information they obtained for the colicin E1 promoter region.

The evidence (1-3) cited in the introduction to this paper firmly establishes that P1 is the true colicin E1 promoter and that it is repressed by lexA binding at the "SOS" box (Fig. 5). We have found some striking homology between P1 and the tet promoter which is shown in Fig. 5. Both cea and tet have been shown by Yang et al. (17) to be strongly dependent on superhelicity for their gene expression, suggesting that the homology may be significant for RNA polymerase-DNA contacts in superhelical DNA. The sequence ATGCTGTATATA A AACC AGT GGTT A TATGTACAGTAT can form a 16 basepair hairpin structure with two wobble G-T basepairs and two bulges. This covers the "SOS" box, the initiation region of P1 and the "Pribnow" box of P2 (Fig. 5). The extrusion of this sequence to form a large imperfect hairpin would be promoted by supercoiling which could act to enhance RNA polymerase binding. The K_d for lexA binding to the lexA operator sequences (Fig. 1) has been estimated to be 20nM (14). Whether this would also be enhanced by supercoiling requires experimental testing. The recent identification of supX mutants with decreased DNA topoisomerase I activity (18,19) and their apparent inability to generate any UV induced mutants (20) suggest the possibility that increased negative supercoiling may prevent the removal of lexA, thereby blocking "SOS" repair and mutagenesis.

The above considerations point out that the natural state of ColE1 DNA is superhelical and structural changes due to supercoiling most likely eli-

minate RNA polymerase binding at P2 and P3 leaving only P1 as the functional promoter for colicin E1 synthesis. We have noted that one can find a stretch of four A residues in each promoter (Fig. 5). For P2 and P3 this sequence homology is spaced exactly 14 nucleotides from the "Pribnow" box. This spacing would be also true for P1 if we utilize the alternative TTTTATG as the "Pribnow" box. For P3 the "Pribnow" box is a perfect consenus sequence for a promoter (9) and may explain the strength of RNA polymerase binding. On the other hand, during replication, lexA has to be removed transiently and transcription could occur from P1. However, if RNA polymerase binding at P3 is favored, under these reduced supercoiling conditions, we could envision active transcription from P3 which would block colicin E1 expression. Alternatively, transcription from P3 could act positively to remove pause sites and enhance the rate of colicin E1 synthesis. Obviously to fully understand cea regulation we will need to explore the above possibilities in detail as well as other considerations arising from the studies of other "SOS" operator regions. Although we have focused some attention in this discussion on the colicin E1 promoter, we view the major contribution of this study to be the developement of new methodology for exploring RNA polymerase-DNA interactions, RNA polymerase-RNA polymerase interactions for promoters in close proximity, and finally the detection of initiation sites by the defined movement of RNA polymerase using ternary complex formation.

ACKNOWLEDGEMENTS

We are grateful to R.T. Sauer, J.L. Suit, S.E. Luria, and W.A. Cramer for their unpublished colicin E1 protein sequence information. This work was supported by Grant CA 17077 from the National Institutes of Health. Peter Chan was supported by a Genetics Training Grant T32 GM07091. We also thank Carolyn Harris for preparation of this manuscript.

REFERENCES

1. Ebina, Y., Kishi, R., Miki, T., Kagamiyama, H., Nakazawa, T., and Nakazawa, A. (1981) Gene 151:119-126.
2. Ebina, Y., Kishi, F., and Nakazawa, A. (1982) J. Bacteriol. 150:1479-1481.
3. Yamada, M., Ebina, Y., Miyata, T., Nakazawa, T. and Nakazawa, A. (1982) Proc. Natl. Acad. Sci. U.S.A. 79:2827-2831.
4. Schmitz, A. and Galas, D.J. (1979) Nucl. Acids Res. 6:111-137.
5. Taylor, W.E. and Burgess, R.R. (1979) Gene 6:331-365.
6. Chan, P.T., Lebowitz, J., and Bastia, D. (1979) Nucl. Acids Res. 7:1247-1262.

7. Maxam, A., and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74:560-564.
8. Shalloway, D., Kleinberger, T. and Livingston, D.M. (1980) Cell 20:411-412.
9. Rosenberg, M. and Court, D. (1979) Annu. Rev. Genet. 13:319-353.
10. Sancar, A., Stachelek, C., Konigsberg, W. and Rupp, W.D. (1980) Proc. Natl. Acad. Sci. U.S.A. 77:2611-2615.
11. Miki, T., Ebina, Y., Kishi, R., and Nakazawa, A. (1981) Nucl. Acids Res. 9:529-543.
12. Horii, T., Ogawa, T., and Ogawa, H. (1981) Cell 23:689-697.
13. Little, J.W., Mount, D.W. and Yanisch-Perron, C.R. (1981) Proc. Natl. Acad. Sci. U.S.A. 78:4199-4203.
14. Brent, R., and Ptashne, M. (1981) Proc. Natl. Acad. Sci. U.S.A. 78:4203-4208.
15. van den Berg, E., Zwetsloot, J., Noordermeer, I., Pannekoek, H., Dekker, B., Dijkema, R. and van Ormondt, H. (1981) Nucl. Acids Res. 9:5623-5643.
16. Levine, A.D. and Rupp, W.D. (1978) in "Microbiology 1978", ed. Schlessinger, D. (Am. Soc. Microbiology) pp 163-166.
17. Yang, H-L., Heller, K., Gellert, M. and Zubay, G. (1979) Proc. Natl. Acad. Sci. U.S.A. 76:3304-3308.
18. Truckis, M. and Depew, R.E. (1981) Proc. Natl. Acad. Sci. U.S.A. 78:2164-2168.
19. Sternglass, R., DiNardo, S., Voelkel, K.A., Nishimura, Y., Hirota, Y., Becherer, K., Zumstein, L. and Wang, J.C. (1981) Proc. Natl. Acad. Sci. U.S.A. 78:2747-2751.
20. Overbye, and Margolin, P. (1981) J. Bacteriol. 146:170-178.
21. Chan, P.T.; Dow Chemical Corp. Midland, Mich. 48640.
22. Lebowitz, Jacob, University of Alabama in Birmingham, University Station, Birmingham, Alabama 35294; to whom reprint request should be addressed.