Mapping of RNA polymerase binding sites in R12 derived plasmids carrying the replication-incompatibility region and the insertion element IS1

Peter T.Chan and Jacob Lebowitz *

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

Received 7 July 1982; Revised and Accepted 11 October 1982

Abstract

Interactions between *Escherichia coli* RNA polymerase holoenzyme and three small plasmid DNAs (pSM1, pSM2, and pSM15) derived from the drug resistant factor R12 have been studied. These plasmids carry the copy number and incompatibility determinants, the origin of DNA replication and the rep gene(s) necessary for plasmid replication. They also contain the insertion element IS1 and the putative finO cistron. Thirteen DNA segments within the largest of the three plasmids (pSM2) were able to form either a binary and/or ternary complex with RNA polymerase. A unique strong binding site was mapped within the left end of IS1. Five binding sites were found within the rep-cop-inc region. Four of these are weak binding sites whereas the fifth does not form a stable binary complex and was detected by ternary complex formation. A strong binding site was located in the putative finO region whereas the remaining six binding sites are located in regions with unidentified genetic functions.

Introduction

A series of small plasmids derived from the drug resistant factor R12 have been isolated in vivo by tetracycline selection (1,2). Three of these plasmids, pSM1, pSM2, and pSM15 are contained within a segment of the R12 extending from 82.7 to 89.3 kilobases, and are of particular interest for the following reasons. These plasmids carry the insertion element IS1 (3), the origin of DNA replication (oriV) (4), repA (5), incompatibility (6,7), and in two of them (pSM1 and pSM2), the reputed "fertility inhibition" (finO) cistron (8,9). The largest of the three is pSM2. In the circular form, pSM1 and pSM15 have deletions to the left and right of the IS1 sequence respectively, compared to pSM2. The genetic map of the plasmid replication functions from the above studies is presented in Fig. 1 with the RNA polymerase binding sites that were determined from transcription and electron microscopy studies (10,11). We have extended the above studies by examining the RNA polymerase-DNA restriction fragment complexes for their binding ability at low (0.05 M KCl) and high (0.2 M KCl) salt

concentrations using nitrocellulose filter assay methodology.

Materials and Methods

Bacterial Strains: Escherichia coli w677/pSM1, 2677/pSM2, and CSH2/pSM15 were kindly provided by Dr. Susan Mickel.

Growth of Bacteria: All strains were grown in Panassay broth (Difco). A 1:100 dilution of an overnight culture was grown at 37°C for 24 hours before harvesting the cells for plasmid DNA purification.

Plasmid DNA purification: An adaptation of the method of Birnboim and Doly (12) for large scale purification of plasmid DNA was used. All plasmid DNAs used have been purified by two successive CsCl density gradient centrifugations.

Restriction endonucleases: The restriction endonucleases HaeII and HinfI were purified according to published methods (13,14). PstI and DdeI were purchased from New England Biolabs. All restriction digestions were carried out in 0.02 M Tris (pH 8), 10 mM MgCl₂ at 37°C.

Binary complex assay: A modification of the method of Jones and Reznikoff (15) was used. The total assay volume used was 0.2 ml in Tris buffer containing 0.04 M Tris, pH 8.0; 0.01 M MgCl₂; 1 mM dithiothreitol; 0.1% autoclaved gelatin; 0.05 M or 0.2 M KCl. In each assay 5 µg of DNA were used. The molecular ratio of RNA polymerase to DNA was 50:1. An extinction coefficient of ϵ_{280} (1%) = 6.5 was used to determine the concentration of purified E. coli RNA polymerase. After a 10 minute incubation of DNA and enzyme at 37°C, heparin (8 mg/ml) was added to a final concentration of 0.4 mg/ml in order to compete out non-specific binding. RNA polymerase-DNA complexes were retained on a Schleicher and Schuell nitrocellulose B-6 filter (pretreated with 0.4 N KOH for 5 minutes), and washed six times with 0.25 ml aliquots of binding buffer minus gelatin. The polymerase bound DNA was extracted from the filter by incubating the filter in 2 ml of 1% SDS at 60° for one hour with occasional shaking. DNA was concentrated by ethanol precipitation in 0.3 M sodium acetate in an ethanol dry ice bath, resuspended in 0.01 M Tris, pH 8; 1 mM EDTA [Na₂]; and then analyzed by gel electrophoresis. The above approach was utilized by us previously (16) to characterize RNA polymerase-DNA interactions for the colicin E1 promoter involved in copy number and incompatibility functions.

Ternary complex assay: A modification of the method of Taylor and Burgess (17) was used. Experimental conditions were the same as in the binary complex assay except for the addition of combinations of two or three ribo-

nucleotide triphosphates (NTPs), obtained from P. L. Biochemicals, to a final concentration of 0.5 mM and KCl to 0.1M in a total assay volume of 0.5 ml. At the end of a 10 minute incubation, the KCl concentration was raised to 0.5M with the addition of 2M KCl and incubation continued for two minutes at 37°C. The six possible combinations of two NTPs (ATP+CTP, ATP+GTP, ATP+UTP, GTP+CTP, GTP+UTP, and CTP+UTP) and the four possible combinations of three NTPs (minus ATP, minus CTP, minus GTP and minus UTP) were used.

In a previous study of a ColE1 DNA promoter (16) an analysis of restriction sites protected by RNA polymerase before and after limited transcription showed that RNA polymerase movement corresponded to the length of the RNA that would be generated with the appropriate set of NTPs until the enzyme stopped at the missing NTP. Subsequent studies on another ColE1 DNA promoter using exonuclease III protection confirmed that transcription was limited by the respective set of NTPs used (Chan and Lebowitz, manuscript submitted for publication). Consequently cross contamination of NTPs does not appear to be a complicating factor in the ternary complex binding assays. In addition chromatographic analysis of recent NTPs obtained from P. L. Biochemicals on PEI cellulose did not reveal cross contamination (D. Wood, unpublished observations) supporting the conclusion drawn above.

Results

Mapping of RNA polymerase binding sites with HinfI, HaeII and PstI digested DNA

The HaeII, HinfI and PstI restriction maps of pSM2, obtained by Rosen et al. (18) were utilized for this study and are presented in Fig. 1. This figure also summarizes the results obtained from binary and ternary RNA polymerase binding assays that will be described below.

Primarily, HinfI digested DNA was used to map RNA polymerase binding sites. The results were further confirmed with restriction fragments generated by HaeII and PstI, especially in the mapping of possible RNA polymerase binding sites in the insertion element IS₁.

RNA polymerase binding to HinfI digested pSM1, pSM2, and pSM15 DNA at 0.05M and 0.2M KCl is shown in Fig. 2. With pSM2, (center panel B), the largest of the three plasmids, a total of nine fragments were shown to bind RNA polymerase. To assess relative binding strength we analyze both the DNA bound to the filter (retentate, lanes iii) and the DNA that passes

Figure 1. A summary of the available data on R12 derived plasmids; pSM1; pSM2; and pSM15 in relationship to the results obtained in this study on the mapping of the RNA polymerase binding sites of these DNAs using filter assay methodology.

Row A - megadalton position for the plasmids on the R12 map

Row B - fractional map units of pSM2, the largest plasmid of the series.

Row C - The PstI restriction map of pSM2 DNA based upon the results of Rosen et al. (18). The subsequent restriction maps are also based on this reference. The deletion in pSM15 produces a subfragment of PstI d which we refer to as d'.

Row D - HaeII restriction map of pSM2 DNA.

Row E - HinfI restriction map of pSM2 DNA.

Rows C-E present the data on the binary complex assay of RNA polymerase-DNA restriction fragments bound to nitrocellulose filters where (▬) refers to fragments with weak binding sites and (■) refers to fragments with strong binding sites (see text for binding criteria). Restriction fragment boxes not closed showed no binary complex formation.

Row F - Location of ternary complexes (▬) on the HinfI map of pSM2.

Row G - Location of the deletions on (1) pSM15 and (2) pSM1. HinfI fragments b and c of pSM2 become respectively subfragments b' in pSM15 and c' in pSM1 and will be referred to as such in the text.

Row H - Summary of the twelve segments of pSM2 DNA containing either binary or ternary complexes.

Row I - Location of four RNA polymerase binding sites (●) which were identified by electron microscopy by Lurz et al. (11).

Row J - The location of the possible coding sequences from the respective DNA sequence references 3, 18 and 22. The location of the *in vitro* transcripts RNA I, II and III from the results of Rosen et al. (19).

Row K - The identified genetic loci from Lurz et al. (11) and the replication protein nomenclature (identified or potential) for the replication-incompatibility region from Rosen et al. (18,19).

Figure 2. RNA polymerase binding to *Hinf*I digested fragments: (A) pSM1, (B) pSM2, and (C) pSM15 DNA in the presence of (a) 0.05M and (b) 0.2M KCl. Lane (i) is *Hinf*I digested DNA, lane (ii) is DNA recovered from filtrate, and lane (iii) is DNA retained by nitrocellulose filter (retentate).

through the filter (filtrate, lanes ii). The low salt condition (0.05M KCl), panel a, allows maximum RNA polymerase binding resistant to a heparin (0.4 mg/ml) challenge and consequently both weak and strong binding sites can not be differentiated. At high salt (0.2 M KCl), panel b, weak binary complexes are dissociated and we can detect the strong binary complexes resistant to a heparin challenge. We evaluated the approximate strengths of the RNA polymerase binding sites by visual inspection of lanes ii and iii in Fig. 2 and estimated the fraction of each fragment bound in 25% ranges (Table 1). Examination of Table 1 immediately reveals that pSM2 *Hinf*I fragments a, b, c and e formed strong binary complexes that were resistant, in varying degrees, to dissociation by 0.2M KCl. Fragments d, g, h, i, k formed weak binary complexes.

In pSM1, (Fig. 2 panel A), both fragments e and i are missing due to deletion. In addition a small segment to the left end of fragment c is deleted. The remaining fragment c' migrated together with fragment d (Fig. 2, panel A). However, since fragment d only contained a weak binding site (Fig. 2, panel B), the strong binding property observed with the c'-d band must be due to c' which retained the strong binding property of fragment c.

The deletion in pSM15 generates *Hinf*I b' and this removes the strong binding property of the *Hinf*I b fragment, but weak binding is still observed for the b' subfragment (Fig. 2, panel C). This suggests a minimum of two RNA polymerase binding sites exist in *Hinf*I b, one strong and one weak.

Lurz et al. (11) and Rosen et al. (19) found two RNA polymerase binding

Table 1. RNA polymerase binding to HinfI restriction fragments of pSM2 and subfragments from deletion plasmids pSM1 and pSM15. Estimates of RNA polymerase binding strengths of the respective fragments by visual inspection at low and high salt concentrations.[†]

Restriction Fragments	0.05M KCl	0.20M KCl
a	++++	++++
b	++++	++
b' (pSM15)	++++	-
c (pSM2, pSM15)	++++	++
c' (pSM1)	+++	++
d	+	-
e	++++	+
f	-	-
g	++	-
h	++++	-
i	++ (weak bands)	-
j	-	-
k	++ (weak bands)	-

[†] The filtrate and retentate for each restriction fragment were visually examined and scored for the percentage bound to the filter as follows: ++++ > 75%; +++ > 50%; ++ > 25%; + < 25% and - no detectable binding.

sites in the HinfI g fragment (Fig. 1) and we conclude that both of these sites are weak based upon the binding data tabulated in Table 1. In summary the binding analysis of HinfI fragments demonstrated that a minimum of ten segments (a, b, b', c, d, e, g, h, i, k) can form a minimum of eleven binding sites (two in fragment g). Of these only a, b, c, and e form strong binary complexes.

Identification of a single strong RNA polymerase binding site in IS1

Examination of the restriction maps of Fig. 1 shows that HinfI c (pSM2) and HinfI c' (pSM1) and the insertion sequence IS1 overlap extensively. This strongly suggests that a strong RNA polymerase binding site is located in the IS1 region. This was confirmed by an RNA polymerase binding study with HaeII digested pSM2 DNA (Fig. 1, row D) in which strong binding was detected with the HaeII e fragment (Fig. 3). No binding occurred to the HaeII f fragment, the middle region of IS1. Binding to the right end of the IS1 sequence was excluded by virtue of experiments with PstI digested pSM2 and pSM15 DNA (Fig. 1, row C). Figure 4 shows that the deletion in pSM15 removed the strong binding in PstI fragment d of pSM2 and generates a subfragment d' without any binding activity. Consequently, the right end of the IS1 sequence which is covered by d' does not contain any RNA polymerase binding sites. The elimination of any RNA polymerase binding to the right end of the IS1 sequence puts the location of the strong binding site

Figure 3. RNA polymerase binding to HaeII digested pSM2 DNA in the presence of (a) 0.05M and (b) 0.2M KCl. Lane (i) is HaeII digested pSM2 DNA, lane (ii) is DNA recovered from filtrate, and lane (iii) is DNA retained by the nitrocellulose filter.

to within 112 basepairs at the left end of IS₁ (from the left end of IS₁ to the HaeII e-f junction).

Binary complexes formed in the replication control region

The results of the binding analysis of PstI fragments of pSM2 and pSM15 (Fig. 4) allowed us to localize the weak binding site in the HinfI b' fragment of pSM15 (Fig. 2, Table 1). This weak site must be located to the right of the PstI d-c junction (Fig. 1). We reach this conclusion from the observation that the PstI d' subfragment (pSM15) has no binding capacity

Figure 4. RNA polymerase binding to PstI digested (A) pSM2 and (B) pSM15 DNA in the presence of (a) 0.05M and (b) 0.2M KCl. Lane (i) is PstI digested DNA, lane (ii) is DNA recovered from filtrate and lane (iii) is DNA retained by nitrocellulose filter.

(Fig. 4) whereas the pSM2 d fragment showed strong binding. Consequently by comparison of the overlap between the PstI d' and HinfI b' fragments (Fig. 1) the weak site in the latter must be located outside of the PstI d fragment to allow for the weak binding observed for the pSM15 b' subfragment. This was supported completely by electron microscopy data (11) as shown in Fig. 1. As indicated previously the strong site in HinfI b and PstI d must be located in the deleted region of pSM15 (Fig. 1), i.e. the putative finO region.

Consequently, in regard to the replication control region, all RNA polymerase-DNA interactions observed within this region were weak. The HinfI g fragment can be seen to be particularly weak considering the fact that it contains two in vitro promoters (11,19), and we still observed more nonbound DNA than bound DNA (Fig. 2, Table 1). The PstI data (Fig. 4) confirms the weak binding for this region. PstI fragment c contains three RNA polymerase binding sites; the one found to the right of the PstI d/c junction in the analysis described previously and the two in the overlap of PstI c with HinfI g. Even with three binding sites PstI fragment c is still totally sensitive to 0.2 M KCl confirming the weakness of RNA polymerase interaction in this region (Fig. 4).

Ternary complex assay

The possibility exists that RNA polymerase binding sites may go undetected by using binary complex formation as the sole criteria for the determination of RNA polymerase-DNA interactions. A very weak binary complex may be stabilized very rapidly by initiation and consequently may only be detected after limited transcription. Initiation can take place if the correct combination of NTPs supports limited RNA synthesis and transcription will stop at the location where the missing nucleotide is required. If transcription has proceeded beyond the third nucleotide, a stable ternary complex can be formed (20). Such ternary complexes of RNA polymerase, DNA and RNA are stable in the presence of 0.5 M KCl (17). This forms the basis of the ternary complex assay. However, one must note that in vitro transcription can initiate within an eight base-pair stretch centered around the origin of transcription, as demonstrated with the phage T7-A3 promoter (21). Therefore more than one combination of NTPs may be able to initiate transcription at the same promoter. From the T7-A3 data (21), the rate of initiation is maximum at the true starting purine and falls off in an exponential fashion in either direction from the true initiating nucleotide. On a qualitative basis the ternary assay is very use-

Figure 5. Ternary complex assay of HinfI digested pSM1 DNA. The unlabeled lane is the HinfI digested DNA. -N is the control in which RNA polymerase and DNA are mixed with no NTPs added followed by raising the salt concentration to 0.5M KCl. -A equals CTP+GTP+UTP; -C equals ATP+GTP+UTP; -G equals ATP+CTP+UTP; -U equals ATP+CTP+GTP; AC equals ATP+CTP; AG equals ATP+GTP; AU equals ATP+UTP; CG equals CTP+GTP; CU equals CTP+UTP; and GU equals GTP+UTP. Each vertical lane represents the capacity of respective restriction fragments to form stable ternary complexes with the above combinations of NTPs.

ful when sequence data are available since we can analyze whether the initiation pattern truly reflects the promoter site in question. Nucleotide sequence data for six HinfI fragments (b', c, d, f, g, h) is available (3,18) and the interactions between RNA polymerase and these fragments will be subjected to an analysis for promoter sequences that can satisfy the pattern of initiations observed in the ternary complex assays.

Figures 5 and 6 show the results of the ternary complex assay in which

Figure 6. Ternary complex assay of HinfI digested pSM15 DNA. The lanes are labeled as Figure 5.

pSM1 and pSM15 HinfI restriction fragments were examined for their ability to form stable RNA polymerase-DNA-RNA complexes initiated by the combination of NTPs shown: -A equals GTP, UTP and CTP; -C equals GTP, ATP and UTP; -G equals ATP, CTP and UTP; AC, AG, AU, CG, CU and GU equals the respective two NTPs combinations. Table 2 presents a summary of the capacity of the respective restriction fragments to form ternary complexes scored visually by examining the intensity of bands in the electrophoretic analysis. The use of pSM1 and pSM15 is sufficient to test for all pSM2 fragments although the latter was examined as well. Consequently for the purposes of easier data presentation it is sufficient to present only the data for the deletion plasmids. The HinfI fragments a, f, g, h, j and k are identical in both pSM1 and pSM15 and consequently two sets of data are available to examine reproducibility. As shown previously pSM2 and pSM15 fragment C overlaps the IS₁ sequence and we established a strong RNA polymerase binding site in IS₁ at its left end. In pSM1 the left end of the IS₁ sequence is joined to part of a small HinfI fragment between fragments k and e (Fig. 1). This generates upon HinfI digestion the fragment we previously designated c' which migrates with fragment d. From the sequence analysis (3) HinfI C contains 102 basepairs of plasmid DNA whereas HinfI C' contains 62 different basepairs of plasmid DNA. However as seen in Figs. 5 and 6 there is little difference in the sets of NTPs that support the formation of ternary complexes between pSM1 fragments cd' and pSM15 fragments c plus d (Table 2). Consequently we would conclude that the strong binding site in the insertion sequence predominates. The question of reproducibility can be examined by comparing identical fragments in pSM1 and pSM15. Except for fragments g and h the reproducibility was good as judged by our visual band intensity assignments. With the same exception as stated above, a comparison of the ternary complex data for pSM2 (not shown) gave good agreement with the data of Figs 5 and 6.

In comparing the ternary complex data between pSM1 and pSM15 we observed that pSM1 fragment HinfI a appeared to have a much stronger initiation for sets of two NTPs than pSM15, which initiated at background levels except for C+U. Since these are identical fragments we conclude that the background for pSM1 is much higher. Examination of the data for pSM2 indicated ternary complex data that was essentially in agreement with pSM15 for fragment a (data not shown). Comparison of pSM1 and pSM2 with regard to fragment b again showed higher backgrounds for pSM1 (data not

Table 2. Summary of the capacity of respective restriction fragments to form ternary transcription complexes initiated by the sets of NTPs shown.

Restriction endonuclease fragments	-A	-C	-G	-U	AC	AG	AU	CG	CU	GU
	GTP	GTP	ATP	ATP	ATP	ATP	ATP	GTP	CTP	GTP
	UTP	ATP	CTP	GTP	CTP	GTP	UTP	GTP	UTP	UTP
pSM1a	+++	+	++++	+++	?	?	?	?	+++	?
pSM15a	+++	+	+++	++	+	-	?	?	+++	-
pSM1b	++	+	++++	+++	++	?	?	?	+++	++++
pSM15b'	+	-	++	++	+	-	-	-	-	-
pSM1c'd	+++	+++	+++	+++	?	?	++	?	++	+++
pSM15c	++	+++	+++	++	-	++	++	-	-	-
pSM15d	++	-	+	-	-	-	-	-	-	-
pSM15e	-	+	+	+	++	++	+++	-	+	++
pSM15f	++	-	+	+	-	-	-	+	-	-
pSM15j	+++	-	+	+	-	-	-	+	-	-
pSM1g	++	+++	+++	++	-	-	+++	-	+	++
pSM15g	+	+	+	+	-	-	++	-	+	+
pSM1h	++	+	+++	+++	+++	++	-	-	-	+
pSM15h	++	+	+	+	+	-	-	-	-	-
pSM15i	-	-	-	-	-	-	-	-	-	-
pSM1j	-	-	-	-	-	-	-	-	-	-
pSM15j	-	-	-	-	-	-	-	-	-	-
pSM1k	-	-	-	-	-	-	-	-	-	-
pSM15k	-	-	++	++	-	-	-	-	-	-

Visual inspection of intensity of the ternary complex bands were scored as follows: - = no band; + = faint band; ++ = moderate intensity; +++ = bright band; ++++ = very intense wide band; ? = possibly resulting from higher background binding.

shown). The reason for this situation is not clear. We have placed a question mark in Table 2 for those cases where high background binding interfered with the visual scoring of bands. However where initiation was strong, e.g. C+U HinfI fragment a, we can readily detect this well above background and the agreement between pSM1 and pSM15 is good. Background effects diminished with smaller fragments due to the much lower amounts of DNA present.

Identification of promoters satisfying the ternary complex data.

Examination of the sequence data for HinfI fragments b', c, d, f, g and h allowed us to identify promoters that would support the pattern of initiation observed with the sets of NTPs used (Table 2). The conclusions are summarized in Fig. 7 and will be discussed below.

The sequence shown for the B' promoter would initiate only with A+C synthesizing AAAC. Addition of U would extend this sequence by a U residue. Consequently, initiation with A+C should be equivalent to initiation with A+C+G(-U) and A+C+U(-G). Although A+C shows a weaker band than -U and -G, the later two bands are identical in intensity supporting the ini-

Figure 7. Nucleotide sequence of possible *in vitro* promoters in the respective restriction fragments that support the formation of ternary complexes with the sets of NTPs shown at the arrows. The nucleotide sequence of *Hinf*I restriction fragments b', d, f, g and h are from Rosen et al. (18) and C1 (pSM2, pSM15) C'1 (pSM1) and C2 (IS1 sequence in all three plasmids) are from Ohtsubo and Ohtsubo (3). The B' promoter is found at nucleotides -4 to +63 ref. 18. The C1, C'1 and C2 promoters are found respectively at nucleotides -48 to +19; -48' to +19 and -1 to +66. The numerical assignments and references for promoters D, F, G1, G2 and H are respectively as follows: 1683-1749 (18); 1068-1142 (18); 353-423 (top coding strand 19); 600-533 (bottom strand 19) and 2276-2342 (top coding strand 19). Arrows indicate initiation sites that fit the ternary complex assay. The underlined portion of the sequence before the initiation site(s) is the potential "Pribnow" box for the promoter. For C2 and F potential ribosome binding sites and initiation codons are indicated by underlining and a combination of overlining and underlining, respectively. For C1, C'a and C2 we repeat the numerical assignments (numbers over dots) given by Ohtsubo and Ohtsubo (3).

tiation assignments. Furthermore, initiation occurs with G+U+C(-A) and the sequence predicts a weak secondary start yielding CUGU. Consequently all observed ternary complexes for the b' fragment can be readily reconciled with the promoter sequence shown in Fig. 7. This promoter sequence is located approximately fifty basepairs preceding the coding region of RepA2 (Fig. 1).

As indicated, fragments c and c' contain the strong RNA polymerase binding site within the left end of IS1. As seen in Table 2 transcription can initiate with all four sets of three NTPs, as well as with A+G, A+U, C+U, and G+U. Initiation also takes place in the absence of GTP. The promoter sequences that best fit this pattern of ternary complex formation are presented in Fig. 7. In pSM15 (pSM2), fragment c contains two potential promoter sequences that we designate C1 and C2. The former contains the junction of plasmid DNA sequence with IS1 and this region shows a "Pribnow" box which would initiate with A+G only if the A residue shown is the true start point or if the adjacent 3'G residue is used. The C2 sequence contains only IS1 residues and the "Pribnow" box starts at +30. The run of five T residues can explain initiation with A+U, C+U and G+U. The strength of the promoter makes it difficult to use visual band intensity differences for a definitive assignment of the starting nucleotide which is most likely the G residue before or after the run of five T residues. In pSM1, subfragment c' contains the C2 sequence but has a different plasmid - IS1 junction sequence which we designate as C'1. Examination of this sequence reveals a potential promoter that would initiate with A+U. Consequently if C'1 is correct we could not distinguish ternary complex formation from C'1 because C2 would also initiate with A+U. Since C'1 does not have an A+G ternary complex and C1 does, we would anticipate a decrease in the A+G ternary complex formation for the c'd pSM1 fragment. There appears to be a significant decrease in ternary complex formation for A+G initiation for c'd relative to c (Figs. 5 and 6), suggesting that our assignment of a promoter in C1 is consistent with the changes in the plasmid - IS1 left junction sequence for pSM1. Weak A+G initiation for the latter probably occurs at the G residue following the run of T residues in the C2 sequence. This supports the start in C2 at the proximal G as shown in Fig. 7 and is consistent with the formation of strong ternary complexes with U alone. This promoter sequence is located just beneath a possible coding sequence for a polypeptide of 90 amino acids long in IS1 (3,22).

In fragment d transcription only initiates well with CTP, GTP and UTP.

A very low level of initiation was observed with ATP, CTP and UTP. Examination of the sequence of this fragment showed a region of sequence containing two possible "Pribnow" boxes with identical initiation specificities, i.e. initiation could occur with G+C+U or with U+C+A (Fig. 7). This promoter region is located approximately 125 basepairs preceding a possible coding sequence near the origin of replication. As shown, this is a weak promoter by binary complex analysis and cannot initiate with any set of two NTPs. Since the distal initiation predicts a possible G+C initiation to form a GCGCC ternary complex and this is not observed, we conclude that the proximal initiation best supports the data for fragment d.

The situation for fragment f is quite interesting since it was shown not to support binary complex formation. However stable ternary complexes are detected with G+U+C (moderately strong) and weaker complexes are formed with A+C+U, A+G+C and G+C. A promoter can be found in the f fragment sequence that meets the observed initiation requirements (Fig. 7). This result is supported by independent observations by S. Kumar and R. Rownd of a promoter in fragment f and its corresponding transcript (personal communication). As shown in the sequence for the fragment f promoter there is a ribosome binding site and start codon and consequently it appears that this promoter could produce an additional protein in the repA1 region (see below).

Fragment g contains two promoters identified by Rosen et al (19), one for the possible repA1 transcript and the other for a small RNA that may be partly responsible for copy number control and incompatibility. These two transcription units overlap each other and transcribe in opposite directions, with repA moving towards the origin of replication (ori) whereas the small RNA moves away from ori (Fig. 1). Rosen et al. (19) have isolated the above RNA transcripts made in vitro and determined the sequence of the 5' end of these RNA molecules. Consequently we can compare ternary complex formation directly with the transcripts made in vitro. In regard to the g fragment we found one DNA sequence that would initiate with G+U and A+U and possibly with U alone (Fig. 7). The latter inference is based upon initiation with C+U which can only occur with the three T residues in the G1 promoter sequence (Fig. 7). Examination of the G2 promoter sequence, the second g fragment promoter, shows that C+U initiation is impossible hence we conclude that a ternary complex is stable but weak with the utilization of the three U residues in G1. The G2 promoter sequence requires A+U+G. These results fit the g fragment promoter assignments

extremely well and are confirmed by the isolation and sequencing of in vitro transcripts by Rosen et al. (19).

Fragment h contains the promoter for RNA III (19) and its behavior is rather unusual. Strong complexes were formed with DNA from pSM1 using A+C, A+G, and as expected A+C+U and A+G+C. However these results were not reproducible in pSM15 and pSM2 (not shown). In particular the A+C and A+G initiated complexes are considerably diminished or virtually undetectable. Although the data for three NTPs appears better we still observed considerably reduced ternary complex formation for pSM15 (pSM2) relative to pSM1 (Table 2). We do not have a satisfactory explanation for the lack of reproducibility with regard to fragment h. The analysis of the transcripts by Rosen et al. (19) shows that the 5' terminal sequence of RNA III is AACAAACAG which fits with A+C initiation. Figure 7 shows the promoter sequence responsible for the RNA III sequence and it is clear that A+G initiation is also possible. Consequently the pSM1 ternary complex data are in complete accord with the in vitro transcript data (19).

Discussion

Although RNA polymerase binding and transcription studies with the replication region of the plasmid R6-5 (which is homologous with R100 and R12) have been previously reported (10), the results presented here provide a more detailed picture of the possible transcription units in the replication-incompatibility region of the R100/R12 plasmids and the IS₁ insertion element.

In regard to the latter our results provide the first direct observation of a strong RNA polymerase binding site at the left end of the IS₁ sequence with initiation occurring at G residue 39. This promoter is C2 of Fig. 7 and the sequences GAGGTG (45 to 50) and the GUG (56-58) provide for ribosome binding and protein initiation respectively. Ohtsubo and Ohtsubo (3) and Johnsrud (22) pointed out this promoter, as well as other potential coding regions, from their respective DNA sequence data. It would appear from our data, which we examined for RNA polymerase binding sites in the middle and right end of the IS₁ sequence, that the left end promoter is the only functional binding site under the in vitro conditions tested. The open reading frame for this site would generate a 91 amino acid polypeptide (nucleotides 56-328) which would be approximately a 10K dalton protein (3,22).

In regard to the replication-incompatibility region Rosen et al. (18)

have identified four open reading frames that can code for the following designated proteins: 11K (repA2); 7K (repA3); 33K (repA1) and 10K (repA4). Row J of Fig. 1 shows the open reading frames and the protein nomenclature is given in Row K. Proteins have been confirmed for repA2 and repA3 (18). In a subsequent study Rosen et al. (19) sequenced in vitro transcripts from superhelical and linear pSM1 and pTRI and found three transcripts designated RNA I, II and III. The latter appears not to be made from superhelical templates. The positions of the transcripts and their direction of synthesis are given in Fig. 1. RNA I, only 91 nucleotides long, is transcribed away from the origin of replication and specifies incompatibility and copy number control functions by mechanisms that most likely involve RNA secondary structure and RNA-RNA interactions with primer RNA (19). RNA II was identified as the mRNA required to produce the repA1 protein (19). The role of RNA III is not clear since it can be deleted without loss of any replicative function (19).

A comparison of the results of Rosen et al. (19) and the electron microscopy data of Lurz et al. (11) with our data show very good agreement (Fig. 1) with the following differences: 1. Rosen et al. (19) did not observe a message from the RNA polymerase binding site in HinfI b' which we found as well as Lurz et al. (11). The position of this site strongly suggests that it codes for repA2 (Cop B). 2. Two additional binding sites were identified with our experiments which were located in HinfI d and f fragments respectively. In both cases, promoter like sequences could be identified that were in agreement with the ternary complex data (Table 2, Fig. 7). Both promoters are close to possible translation initiation signals. In regard to HinfI f, this promoter was only detected as a ternary complex. Since fragment f is completely located within the coding sequence of the repA1 protein, it is surprising to find an in vitro promoter located in this region.

The focus of this study has centered on the regions of pSM2 that have been sequenced. It is clear that considerably more information is needed to answer questions regarding other RNA polymerase binding regions in order to determine both the number and strength of the respective promoters of those fragments.

* To whom reprints should be requested.

Acknowledgements

This work was partially supported by grant CA 17077 from the National Institutes of Health. Peter Chan was supported by Genetics Training Grant T32 GM 07091. It is a pleasure to thank Carolyn Harris for the preparation of the manuscript.

References

1. Mickel, S., and Bauer, W. (1976) J. Bacteriol. 127:644-655.
2. Mickel, S., Ohtsubo, E., and Bauer, W. (1977) Gene 2:193-210.
3. Ohtsubo, H., and Ohtsubo, E. (1978) Proc. Natl. Acad. Sci. USA 75:615-619.
4. Ohtsubo, E., Feingold, J., Ohtsubo, H., Mickel, S., and Bauer, W. (1977) Plasmid 1:8-18.
5. Andre, I., Slocombe, P.M., Cabella, I., Timmis, J.K., Lurz, R., Burkardt, H.J., and Timmis, K.N. (1979) Mol. Gen. Genet. 168:1-25.
6. Taylor, D.P., and Cohen, S.N. (1979) J. Bacteriol. 137:92-104.
7. Miki, T., Easton, A.M., and Rownd, R.H. (1980) J. Bacteriol. 141:87-99.
8. Timmis, K.N., Andres, I., and Achtman, M. (1978) Proc. Natl. Acad. Sci. USA 75:5836-5840.
9. Dempsey, W.B., and McIntire, S.A. (1979) Mol. Gen. Genet. 176:319-334.
10. Chan, P.T., and Lebowitz, J. (1980) Abstracts of the annual meetings of the American Society for Microbiology p.109.
11. Lurz, R., Danbara, H., Ruckert, B., and Timmis, K.N. (1981) Mol. Gen. Genet. 183:490-496.
12. Birnboim, H.C., and Doly, J. (1979) Nucl. Acids Res. 7:1513-1523.
13. Roberts, R.J., Breitmeyer, J.B., Tabachnik, N.F., and Myers, P.A. (1975) J. Mol. Biol. 91:121-123.
14. George, J., and Chirikjian, J.G. (1978) Nucl. Acids Res. 5:2223-2232.
15. Jones, B.B., and Reznikoff, W.S. (1977) J. Bacteriol. 132:270-281.
16. Chan, P.T., Lebowitz, J., and Bastia, D. (1979) Nucl. Acids Res. 7:1247-1262.
17. Taylor, W.E., and Burgess, R.R. (1979) Gene 6:331-365.
18. Rosen, J., Ryder, T., Inokuchi, H., Ohtsubo, H., and Ohtsubo, E. (1980) Mol. Gen. Genet. 179:527-534.
19. Rosen, J., Ryder, T., Ohtsubo, H. and Ohtsubo, E. (1981) Nature 290:794-797.
20. Sylvester, J.E., and Cashel, M. (1980) Biochem. 9:1069-1074.
21. Smagowicz, W.J., and Scheit, K.H. (1978) Nucl. Acids Res. 5:1919-1932.
22. Johnsrud, L. (1979) Mol. Gen. Genet. 169:213-218.