Nucleic Acids Research, Vol. 20, No. 11 2777–2784

Functionally distinct RNA polymerase binding sites in the phage Mu *mom* promoter region

Virginia Balke, Valakunja Nagaraja+, Tracy Gindlesperger and Stanley Hattman*
University of Rochester, Department of Biology, Rochester, NY 14627, USA

Received February 18, 1992; Revised and Accepted April 27, 1992

ABSTRACT

Transcription of the phage Mu *com/mom* operon is trans-activated by another phage gene product, C, a site-specific DNA binding protein. To gain insight into the mechanism by which C activates transcription, we carried out footprinting analyses of *Escherichia coli* RNA polymerase (= RNAP) binding to various *com-lacZ* fusion plasmids. $KMnO_4$-sensitive sites (diagnostic of the melted regions in open-complexes) and DNase I-sensitive sites were located by primer-extension analysis. The results are summarized as follows: (i) In vivo, in the absence of C, RNAP bound in the wild-type (wt) promoter region at a site designated P2; in vitro DNase I-footprinting showed that P2 extends from −74 to −24 with respect to transcription initiation. This overlaps a known strong C-binding site (at −35 to −54). RNAP bound at P2 appeared to be in an open-complex, as evidenced by the presence of $KMnO_4$-hypersensitive sites. (ii) In contrast, when C was present in vivo, RNAP bound in the wt promoter region at a different site, designated P1, located downstream and partially overlapping P2. RNAP bound at P1 also appeared to be in an open-complex, as evidenced by the presence of $KMnO_4$-hypersensitive sites. (iii) Two C-independent mutants, which initiate transcription at the same position as the wt, were also analyzed. In vivo, in the absence of C, RNAP bound mutant *tin7* (contains a T to G substitution at −14) predominantly at P1; in vitro DNase I-footprinting showed that P1 extends from −56 to +21. With mutant *tin6* (a 63 base-pair deletion removing P2, as well as part of P1 and the C-binding site from −35 to −54), RNAP bound to P1 independent of C. We conclude that P1 is the 'functional' RNAP binding site for *mom*-transcription initiation, and that C activates transcription by promoting binding at P1, while blocking binding at P2.

INTRODUCTION

Studies on the expression of bacterial operons have shown that transcription is regulated by a variety of intricate mechanisms. The most thoroughly studied of these has been the interactions of accessory proteins and RNA polymerase (RNAP) with nucleotide (nt) sequences in the promoter region. Recent studies have shown that multiple RNAP- and regulatory protein-binding sites can also play a role in regulating transcription (1–4).

In phage Mu two overlapping genes, *com* and *mom*, are in an operon for which multifaceted controls have evolved to tightly regulate the production of the *mom* gene product (5). At the transcriptional level, methylation of adenine residues in three GATC sites upstream of the *mom* promoter by the *E. coli* DNA-[N^6-adenine]-methyltransferase (Dam) is required for transcription (6–9). This methylation appears to prevent binding of the *E. coli* OxyR protein, which has a recognition site in this region and acts as a repressor when Mu DNA is not methylated at these GATC's (10). Even when these sites are methylated a phage-encoded transcriptional activator protein, C, is required (11–13); C activates transcription at three other late promoters as well. From an alignment of these four late promoter sequences, a site important for C binding has been postulated (13). In vitro DNase I-footprinting of C showed multiple binding sites in the *mom* promoter region, including one site overlapping the −10 hexamer and spacer (−23 to −5), and one site 5′ to and overlapping the −35 hexamer (−52 to −33) (V.Nagaraja, T.Gindlesperger, V.Balke, and S.Hattman, in preparation); hereafter, these sites will be referred to as the −10 and −35 sites. Using MPE·Fe(II) footprinting other workers observed only the −35 binding site (14).

To help elucidate the mechanism by which C activates transcription, we carried out footprinting analyses of *E. coli* RNAP binding to various *com-lacZ* fusion plasmids in the presence/absence of C. $KMnO_4$-hypersensitive sites, diagnostic of the melted regions in open-complexes (15), and DNase I-sensitive sites were located by primer-extension analysis. The results revealed the presence of two functionally distinct RNAP-binding sites within the *mom* promoter region. In the absence of C, RNAP bound predominantly at an upstream site (designated P2) which overlaps the −35 binding site for C. Although RNAP appeared to form an open-complex at P2, transcription in vivo was severely reduced. When C was present in vivo, RNAP bound to a different site, designated P1, located downstream from and partially overlapping with P2. The results with two C-independent mutants support the notion that P1 is the functional binding site for *mom* transcription initiation.

* To whom correspondence should be addressed

+ Present address: Centre for Genetic Engineering, Indian Institute of Science, Bangalore-560012, India

MATERIALS AND METHODS

Materials and general methods

Enzymes and linkers used for DNA cleavage and cloning were purchased from New England Biolabs (NEB) or Bethesda Research Laboratories (BRL). AMV reverse transcriptase (RTase) (200 U/μl) and DNA sequencing kits were from United States Biochemical (USB). RNase-inhibitor was from NEB and DNase I was from Worthington Biochemical Corp. The Klenow fragment of *DNA polymerase I (polIK) was from both* NEB and Pharmacia. *E. coli* RNAP holoenzyme was a generous gift from D. Hinkle and it was also purchased from Pharmacia.

Media and β-galactosidase assays were as described (16). Standard protocols were used for plasmid isolation, restriction digestions, ligation, plasmid transformation, gel electrophoresis and electroelution of fragments from agarose gels (17). For DNA ligation, fragments were isolated from gels using Geneclean™ (Bio 101). Dideoxy-sequencing of DNA was with either [α-^{35}S-NTP] (10 mCi/ml; >1000 Ci/mmol, Amersham) or [^{32}P]-end labeled primers. Primers were end-labeled using [γ-^{32}P]-ATP (10 mCi/ml; >5000 Ci/mmol, Amersham) and T4 polynucleotide kinase (NEB) according to the supplier's instructions.

Hepes, DTT, PEG 8000, heparin and other chemicals were from Sigma Chemical Co. Acrylamide solution for sequencing gels was from National Diagnostics. Rifampicin and yeast tRNA were from Sigma Chemical Co. The *lac* primer (NEB) is complementary to the non-transcribed (top) strand; primer 2, complementary to a vector sequence on the transcribed (bottom) strand, was synthesized at the University of Rochester Sequencing Facility. Prior to use the primers were passed through a PD-10 Sephadex G-25M column (Pharmacia).

Bacterial strains and plasmids

E. coli strain JM83 Δ(*pro-lac*) *rpsL thi ara* Φ80 d*lacZ* ΔM15 was purchased from BRL; strain LL306 Δ(*pro-lac*) *nalA recA supE44 thi* was from L.Lindahl (18). *E. coli* GM1900, an F′ *lacI*^{q} derivative of the prototroph, W3110, was obtained from M.Marinus. Expression vector, pKK223-3, was from Pharmacia and pUC18 was from BRL. Plasmid pACYC184 (19) was provided by L.Lindahl and pKN50 (20) was a generous gift from M.M.Howe; the latter contains the phage Mu C gene and a portion of the *lys* gene. All plasmids were isolated by the alkaline lysis method (21) and purified by CsCl-EtBr centrifugation (17).

Plasmid pKN50 was linearized with *Hpa*I; after the addition of phosphorylated *Bam*HI linker, the DNA was cleaved with either *Bam*HI or *Pvu*I. In the former case, the fragment was cloned into the *Bam*HI site of pUC19 to produce pVN9. In the latter instance, the resulting 3′ ends were made blunt using T4 DNA polymerase. Following *Bam*HI digestion and ligation of *Xba*I linker to the blunt end, the DNA was cleaved with *Xba*I and the *C* gene-containing 800 base pair (bp)-fragment was purified after gel electrophoresis. It was then cloned into the *Bam*HI and *Xba*I sites of pUC18. This plasmid was designated pVN8. Plasmid pVN184 was constructed by cloning the promoterless *C* gene-containing *Bam*HI fragment from pVN9 into the corresponding site (within the *tet* gene) in pACYC184; this placed transcription of *C* under the *tet* promoter.

The *mom-lacZ* fusion plasmid, pMLF-2 (12), was propagated in a *dam*$^{-}$ host. The plasmid DNA was isolated and cleaved at the unique *Bcl*I site within *com*. The 3′ ends were removed by S1 nuclease digestion. After cleavage with EcoRI, the 223 bp fragment (containing the *mom* operon regulatory region and the 5′ end of *com*) was gel purified and ligated into the *Eco*RI/*Sma*I sites of pNM480 (21). This construct (designated pLW4) produces a *com-lacZ* protein-fusion only in the presence of C (unpublished observation).

Isolation of *tin* mutants

Approximately 4×10^{8} cells of *E. coli* JM83 [pLW4] were spread on MacConkey-lactose-ampicillin agar plates and red papillae were purified by streaking. Plasmid DNA from minipreps (21) was analyzed by restriction nuclease digestion to screen for detectable insertions or deletions. Nucleotide sequence anaylsis was used to define the precise mutational alteration.

Construction of single-copy *com-lacZ* fusion λ prophages

The procedures used were essentially those described by Simons et al (24). In brief, the 223 bp *Eco*RI/*Bam*HI fragments from pLW4, pLW4*tin6* and pLW4*tin7*, respectively, were cloned into the corresponding sites of pRS552 and transformed into *E. coli* JM83 (selecting for Km^{R} clones). The resulting plasmids were designated pRSW, pRST6 and pRST7, respectively. λRS45 plate lysates were prepared on these strains and the phage used to transduce *E. coli* LL306 to Km^{R}. Transductant clones were screened to distinguish between monolysogens and dilysogens; wt, *tin6* and *tin7* monolysogens were designated λRS::RSW, λRS::RST6 and λRS::RST7, respectively.

Primer extension of *mom* mRNA

E. coli strain LL306 transformed with either pLW4 ± pVN184, pLW4*tin6* or pLW4*tin7* was grown in 25 ml of L-broth to mid-log phase. Total cellular nucleic acids were isolated (25) and 4 μl (10 μg) of nucleic acid was added to 2 μl of RTase buffer (250 mM Tris HCl, pH 8.3, 500 mM KCl, 50 mM $MgCl_2$, 5 mM DTT, 5 mg/ml bovine serum albumin). After the addition of 1μl (750,000 cpm) of end-labeled *lac* primer and 1 μl of RNase Inhibitor, the primer was allowed to hybridize for 2 hr at 42°C. Then 1 μl (200 U) of AMV RTase and 1 μl of dNTP mix (20 mM) were added and the reaction allowed to proceed for 1 hr at 37°C. After the addition of 2 μl of Stop-Dye Mix (0.1% xylene cyanol, 0.1% bromophenol blue, 10 mM EDTA, 95% deionized formamide), the samples were diluted 1:4, heated at 95°C for 2 min and subjected to electrophoresis in a 6% thick polyacrylamide sequencing gel.

In vitro transcription

The 223 bp EcoRI-BamHI fragment from pLW4*tin7* was isolated by electroelution from an agarose gel. The transcription assay (26) was as follows. Approximately 50 ng of fragments were brought to a final volume of 25 μl with the addition of 5 μl of 5×transcription buffer (600 mM KCl, 50 mM $MgCl_2$, 0.5 mM DTT, 0.5 mM EDTA, 50 mM Tris HCl, pH 8.0, 25% glycerol), 2.5 μl NTP mixture and 5 μCi or either [γ-^{32}P]-ATP or [γ-^{32}P]-GTP (10 mCi/ml; >1000 Ci/mmol, Amersham). The final concentration of the NTP corresponding to the radioactive nt was 0.2 mM, while the three other nts were at 3 mM. The mixtures were incubated for 3 min at 37°C and then 1 μl of RNAP was added. After 5 min the concentration of the more dilute radioactive nt was increased to 3 mM by the addition of 1 μl of 7.5 mM NTP. The reaction was continued for an additional 10 min, then 5 μl of Stop-Dye Mix was added and the samples were subjected to electrophoresis in a 6% polyacrylamide gel.

$KMnO_4$-Footprinting

Slight modifications of published protocols (15) for in vivo and in vitro treatment with $KMnO_4$ were followed. Cultures of LL306 transformed with the appropriate plasmid(s) were grown in M9-glucose-ampicillin medium supplemented with 20 mg/ml casamino acids. When the culture reached mid-log phase, it was divided into two 10 ml aliquots and one was treated with 200 μg/ml rifampicin for 5 min (to trap RNAP in the open-complex). Then both samples were treated with $KMnO_4$ for 4 min and poured into 30 ml Corex tubes containing crushed ice. Plasmid DNA was isolated as described (17) except that the samples were extracted once with phenol prior to precipitation with 2-propanol. For in vitro footprinting of RNAP, 2 μg of plasmid DNA were brought to 16 μl in Hepes transcription buffer (10 mM Hepes, pH 7.6, 0.1 mM EDTA, 10 mM $MgCl_2$, 4 mM $CaCl_2$, 1 mM DTT, 3% glycerol, 2% PEG 8000). After 3 min incubation at 37°C, 1 μl RNAP (previously diluted 1:10 to 8 μM in 1×Hepes buffer) was added. RNAP was allowed to form complexes for 3 min at 37°C prior to the addition of 4 μl of 100 mM $KMnO_4$. After 4 min incubation at 37°C, 3 μl of β-mercaptoethanol (14.7 M) was added to quench the $KMnO_4$. The samples were passed through a Sephadex G50−80 spin column equilibrated with water and the DNA precipitated by the addition of 4 μl 3 M sodium acetate, 3 μl 5 mg/ml carrier yeast tRNA and 100 μl 95% ethanol. The DNA was resuspended in 75 μl of doubly distilled H_2O prior to the primer extension reaction (see below).

DNase I-footprinting

Two μg (0.34 pmole) of plasmid pLW4 or pLW4*tin7* was incubated with RNAP (12 to 24 pmol) in footprinting buffer (10 mM Hepes, pH 7.6, 0.1 mM EDTA, pH 8.0, 10 mM $MgCl_2$, 4 mM $CaCl_2$, 1 mM DTT, 3% glycerol, 60 mM KCl and 2% PEG 8000). After 5 min at 22°C, 1 μl of DNase I (1 mg/ml in footprinting buffer) was added to the reaction mix for 1 min at 22°C. The reactions were stopped by the addition of 20 ml Stop Buffer (150 mM NaCl, 15 mM EDTA, 50 mM Tris-HCl, pH 7.5, 0.3% SDS). The samples were extracted successively with phenol, phenol/chloroform, chloroform and then precipitated with ethanol and washed. The DNA was resuspended in 70 μl of water and divided into two aliquots. Each aliquot was incubated with a separate end-labeled primer complementary to the top or bottom strand. After alkali denaturation at 80°C, the primer was annealed at 50°C in 0.05 M Tris (pH 7.2), 0.01 M $MgSO_4$, 0.2 mM DTT. After the addition of 5 mM dNTP, 1 μl of polIK (1 unit) was added and primer extension reactions were carried out (see below).

Primer extension

The *lac* primer, which hybridizes within the *lac* gene of the *com-lacZ* fusion, was used to probe modification/cleavage of the top strand, and primer 2 was used to probe the bottom strand. Primer extension reactions were as described (27, 28). Hybridizations were carried out at 50°C when using the *lac* primer and 45°C for primer 2. After precipitation, the extension products were resuspended into 4 μl of Stop-Dye Mix and heated for 2 min at 95°C prior to gel electrophoresis.

RESULTS

Isolation of *tin* mutants

To facilitate studies on the interaction of RNAP with the *com/mom*-operon promoter region, we constructed the *com-lacZ* fusion plasmid, pLW4. Cells harboring pLW4 alone do not contain any β-galactosidase activity due to the absence of Mu C protein, which is required for transcription. This property allowed us to screen for the production of mutations in pLW4 (induced or spontaneous) that led to C-independent transcription and production of the *com-lacZ* fusion protein, as evidenced by the appearance of red papillae over bacterial lawns plated on MacConkey-lactose agar. Plasmid DNA from spontaneous mutants was taken for restriction nuclease and nt sequence analysis; two mutations are described in detail. One (*tin7*) is a T to G transition (corresponding to nt 985 from the right end of the Mu genome) within a T_6-run at −14 with respect to the transcription start site. Another (*tin6*), a 63 nt deletion from −32 to −94, has removed both the −35 binding site for C and the −35 hexamer of the *mom* promoter (Fig. 1). The formation of *tin6* and *tin7* appears to be highly favorable since they both appeared in other mutant screens (S.Hattman and J.Ives, unpublished; V.Balke, unpublished).

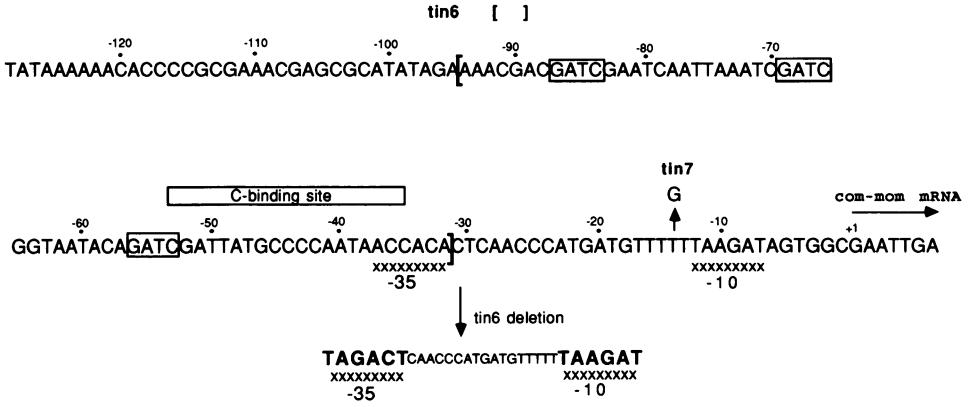

Fig. 1. Mu *mom* promoter region nucleotide sequence; this corresponds to the top strand referred to in Fig.6. The −35 and −10 hexamers are underscored with x's. The *tin7* mutation is a T to G substitution at position −14 with respect to the transcription start site. The *tin6* mutation is a 63 nt deletion (within brackets); the junction created by the deletion forms an improved −35 sequence and spacer length (17 nt), as shown at the bottom. The C-binding site at −35 is denoted by the open bar; Dam methylation sites, GATC, are enclosed in rectangles.

β-galactosidase production by various *com-lacZ* prophage fusions

The level of β-galactosidase was determined for cells containing various forms of pLW4 present as single-copy *com-lacZ* prophage fusions. These constructs were made in order to avoid possible differences in copy-number and maintaining two different plasmids in the same cell. β-galactosidase levels in lysogens λRS::RSW (wt), λRS::RST6 (*tin6*) and λRS::RST7 (*tin7*) with or without a C-producing plasmid, pVN8, were measured. Comparison of the enzyme levels (Table 1) showed that, in the absence of C, the *tin6* and *tin7* lysogens produced 7 and 2-fold more enzyme, respectively, than the wt alone. When the C-producing plasmid, pVN8, was present both the wt and *tin7* prophages were transactivated to similar levels; in contrast, *tin6* was not transactivated, which is consistent with the fact that the −35 C-binding site is deleted in this mutant. Although *tin7* was expressed in the absence of C (albeit weakly compared to *tin6*), it was transactivated to a level about 50% higher than *tin6* or the C-activated wt.

Transcription start site in the *tin* mutants

Previous experiments with S1 nuclease mapping led to different assignments of the transcription start site in the *mom* promoter region (29, 30). More recent studies (14) indicated the start site to be at one of two possible nts, but different from either of those previously reported. Therefore, an alternative method to determine the start site was undertaken, namely, primer extension with RTase. To study transcription from the wt promoter, cells were transformed with both pLW4 and the compatible C-producing plasmid, pVN184. RNA was isolated from these cells, as well as from cells containing pLW4 alone. After hybridization of end-labeled *lac* primer and extension with RTase, the products were analyzed by gel electrophoresis (Fig. 2). In the absence of C, pLW4 produced barely detectable levels of *com-lacZ* mRNA (lane 2). However, when an excess of RNA was analyzed, we did observe some *com-lacZ* mRNA transcripts (lane 1), some of which appear to have initiated at the proposed wt promoter (band corresponding to nt 971). In contrast, when C was present extensive transcription of pLW4 occurred, with the predominant band occurring at position 971 (lane 3). Analysis of mRNA from cells containing either pLW4*tin6* (lane 4) or pLW4*tin7* (lane 5) showed that they also initiated at nt 971. [This position of the start-site is several nt upstream from the one we deduced by S1 nuclease protection analysis (29). Examination of the original autoradiograms revealed that we had discounted the importance of a 'smear' of longer fragments that, in retrospect, must have reflected continued S1 nuclease digestion of the AT-rich end.] The relative levels of transcripts produced by the three promoters correlates well with those of β-galactosidase activity (Table I). These results are consistent with the proposed −35 and −10 promoter hexamer sequences (Fig. 1) and they show that a new transcription start site was not generated in either mutant.

In order to verify that transcription starts at nt 971 (a G), in vitro transcription assays were done using either [γ-^{32}P] ATP or [γ-^{32}P] GTP. Under these conditions, the only transcripts labeled should be those that have incorporated the [γ-^{32}P] label at the 5′ end. RNAP was allowed to transcribe off a *tin7* fragment in the presence of either [γ-^{32}P] ATP or [γ-^{32}P] GTP; we observed that only [γ-^{32}P] GTP produced labeled *mom* transcript (data not shown). This confirms that the start site is G971. Hereafter, all nt positions will be numbered by their position relative to the transcriptional start site, which is defined as +1.

In vivo $KMnO_4$-footprinting of open-complexes

The above results indicated that transcription of the wt promoter was strongly reduced when C was not present. This is not surprising in view of the poor homology to consensus and the suboptimal spacing between the −10 and −35 hexamers. Therefore, it was of considerable interest to study the binding

Table I. Production of β-galactosidase activity in *E. coli* LL306 by various *single-copy com-lacZ* prophage fusions[a]

Strain[b]	β-galactosidase activity[c] (Miller units)
(λRS::RSW)	24 ± 6
(λRS::RSW)[pVN8]	137 ± 20
(λRS::RST6)	163 ± 26
(λRS::RST6)[pVN8]	156 ± 32
(λRS::RST7)	42 ± 6
(λRS::RST7)[pVN8]	222 ± 34

[a] Overnight cultures grown in LB + kanamycin ± ampicillin were diluted 1:50 into fresh medium at 37°C and grown to log phase. Aliquots were taken for assay of β-galactosidase and optical density; corrections for the contribution of LB to the optical density were made in each case.
[b] Prophages are denoted as follows; W = wt; T6 = *tin6*; T7 = *tin7*.
[c] The values represent the mean ± SD from three independent experiments in which triplicate assays were done for each culture.

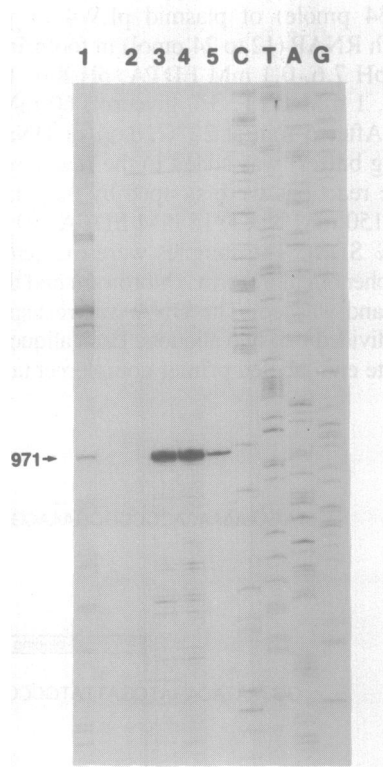

Fig. 2. In vivo analysis of *mom*-transcription initiation by primer extension with RTase. RNA from various plasmid-containing cells was isolated and the 5′ end of the *mom*-specific transcripts analyzed. Lanes 1 & 2, pLW4 alone; lane 3, pLW4 and pVN184; lane 4, pLW4*tin6*; lane 5, pLW4*tin7*. Lane 1 contained a 20-fold higher amount of RNA than analyzed in lanes 2−5. C, T, A, G refer to the top strand sequence, derived from reactions with the complementary dideoxynucleotides.

of RNAP to the wt and *tin* promoters, as well as the effect of C. This was accomplished through an in vivo footprinting analysis with $KMnO_4$. Because this agent preferentially attacks pyrimidines in single-stranded DNA, it is a useful probe for open-complexes (31).

Cells transformed with pLW4wt, pLW4*tin6* or pLW4*tin7* were treated with $KMnO_4$ in the absence or presence of rifampicin [in order to trap RNAP in an open-complex (15)]. Plasmid DNA was isolated and, after hybridization and extension of end-labeled primers, the products were analyzed by polyacrylamide gel electrophoresis to determine the sites of modification on each of the two strands (results for the top strand are shown in Fig. 3; + RNAP denotes addition of rifampicin). In the presence of rifampicin, there were sites of strongly enhanced $KMnO_4$ sensitivity (hypersites) at positions −47 and −41 on the top strand of pLW4wt (compare lanes 1 and 2). Because the appearance of these hypersites was dependent on the presence of rifampicin, we conclude that RNAP binding was required. The location of the hypersites is suggestive of RNAP binding at a site upstream from the proposed −10 and −35 hexamers (Fig. 1). Strong support for this was obtained from in vitro analyses presented below; we designate this upstream binding site as P2. In vivo binding of RNAP to the wt DNA in the presence of C was investigated by probing cells that had been transformed with both pLW4 and the C-producing plasmid, pVN184. In this instance, treatment with $KMnO_4$ produced hypersites at −4 and −3, as well as a weak enhancement at +5 (lane 4). Modification at these positions is consistent with RNAP binding to a site, designated P1, corresponding to an active promoter. Support for this came from analyzing the in vivo binding of RNAP to plasmids *pLW4tin6* and *pLW4tin7*. For example, in the absence of C, hypersites corresponding to P1 were produced on tin7 at −4, −3 and +5 (lane 6); in addition, some $KMnO_4$-modification was still seen at −47 and −41. These results indicate that there may be competition for RNAP binding at P1 and P2. In the presence of C, RNAP bound predominantly to P1 (lane 8). With *tin6* (lanes 10 and 12), only P1 hypersites were observed. This result was expected since P2 and the −35 C-binding site are deleted in this mutant. It should be noted that modification of the +5 site was reduced in the presence of C (compare lanes 6 and 8, and lanes 10 and 12).

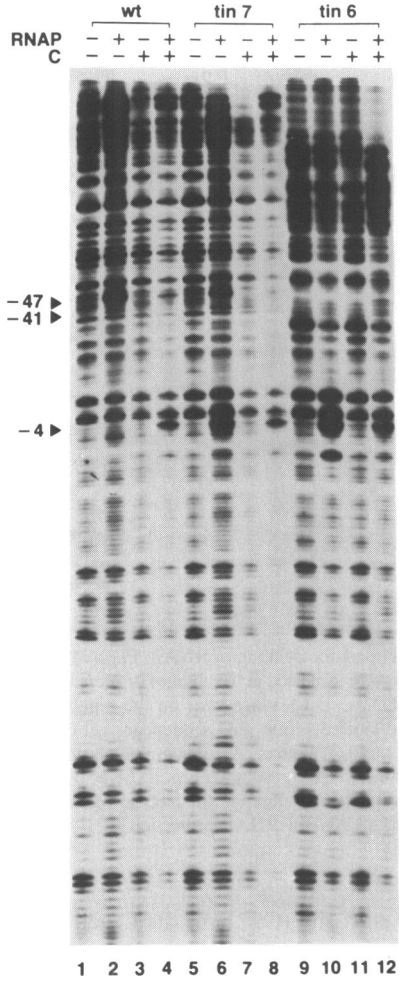

Fig. 3. In vivo analysis of $KMnO_4$-sensitive sites (top strand) on wt and *tin* plasmid DNAs. DNA was isolated from pLW4-containing cells [without or with plasmid pVN184 to provide C protein (indicated with a + on the appropriate line)] after treatment with $KMnO_4$ and analyzed by primer-extension sequencing. In this experiment, rifampicin had to be added in order to trap RNAP in the open-complex; + RNAP and − RNAP denote presence and absence of the drug, respectively.

In vitro $KMnO_4$ and DNase I-footprinting of RNAP

In order to confirm the existence of the two RNAP-binding sites, P1 and P2, an in vitro $KMnO_4$-footprinting analysis was carried out. RNAP was bound to supercoiled wt and *tin7* plasmid DNAs in the absence of added nucleotides, and the complexes were probed with $KMnO_4$. On the wt promoter, hypersites were observed at −47 on the top strand, and at −55, −57 and −62 on the bottom strand (Fig. 4; compare the lanes marked − and + denoting absence and presence of RNAP, respectively). In contrast, with pLW4*tin7*, hypersites were observed at −4, −3 and +5 on the top strand, and at −10 and −11 on the bottom

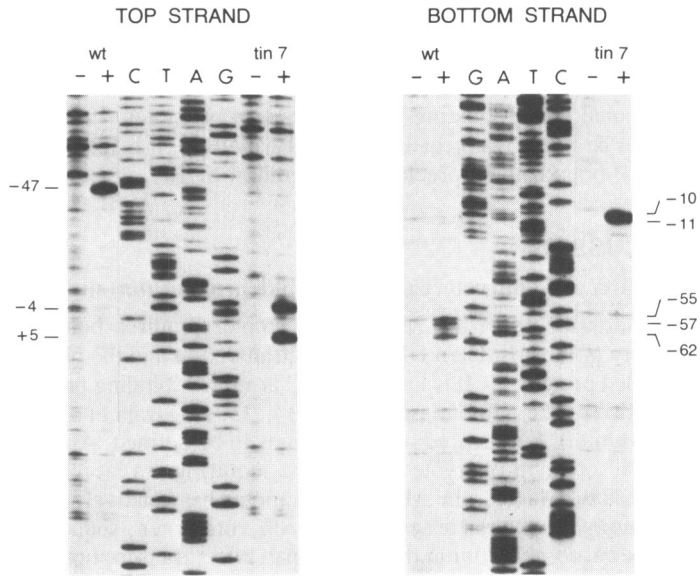

Fig. 4. In vitro analysis of $KMnO_4$-sensitive sites on the top and bottom strands of wt and *tin7* plasmid DNAs. The symbols + and − denote addition and omission of RNAP, respectively. Sequencing lanes C, A, G and T are shown.

Fig. 5. Summary of footprinting with $KMnO_4$ and RNAP in the *mom* promoter region. $KMnO_4$-modified sites are indicated with one caret (moderate sensitivity) or two carets (high sensitivity). The −10 and −35 sequences in P1 (top strand) are enclosed in rectangles; putative −10 and −35 sequences in P2 are overlined (top strand) or underlined (bottom strand).

strand; these sites are characteristic of open-complexes at known functional promoters. A composite of the in vivo and in vitro $KMnO_4$-sensitive sites is shown in Fig. 5.

To further investigate RNAP binding to the *mom*-promoter region, we carried out an in vitro DNase I-footprinting analysis (32). RNAP was bound to wt and *tin7* plasmid DNAs and the complexes were incubated with DNase I; cleavage sites were identified by extension of end-labeled primer and gel electrophoresis. It is clear from the results (Fig. 6) that different patterns of DNase I cleavage were observed with the two substrates. The DNase I-footprint corresponding to P2 on the wt extends from −74 to −24 (on the top strand); the P1 region on *tin7* spans from −56 to +21. Moreover, in addition to protection against DNase I action, there was also some enhancement of cleavage, particularly at several sites on *tin7* DNA. It is interesting to note that, even in the absence of added RNAP, the wt and tin7 DNAs exhibited different sensitivities to DNase I cleavage in the 'spacer' region between the −10 and −35 sequences. In fact, we have also observed differences to other chemical modifications in this region (unpublished).

These results indicate that in the absence of the transcriptional activator protein C, RNAs capable of binding in vitro to the wt *mom* promoter region. Based on its sensitivity to $KMnO_4$ it appears that RNAP formed an open-complex, or one that had some conformational distortion allowing access to this agent. However, binding at this site, designated P2, is different from the one on *tin7* DNA, designated P1, which appears to be the functional RNAP-binding site for the *mom* promoter. Lastly, it should be noted that because P2 and P1 overlap, it is likely that RNAP bound in P2 precludes recognition of P1.

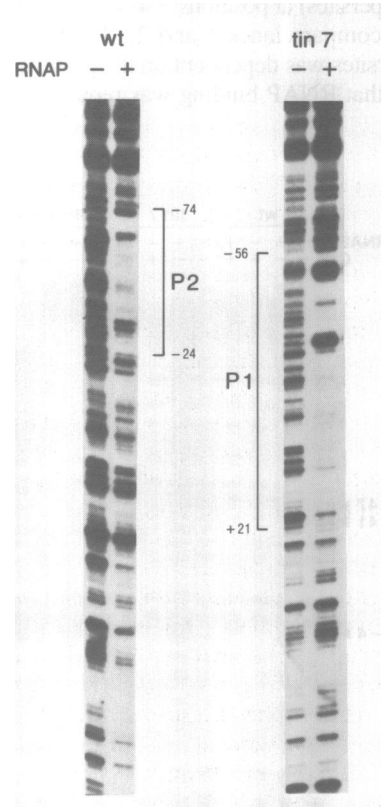

Fig. 6. In vitro DNase I-footprinting of RNAP. Plasmids pLW4 and pLW4*tin7* (2 μg; 0.34 pmole) were incubated in the absence or presence of RNAP (12 pmole) and treated with DNase I. Cleavage sites on the top strand were mapped by extension with end-labeled primer and sequencing gel electrophoresis; control sequencing lanes are not shown. Assuming the plasmids contain only four promoters and that the RNAP preparation was 100% active, then the RNAP:*mom* promoter ratio was 9; both assumptions are likely to be conservative. The vertical brackets denote the regions of RNAP protection.

DISCUSSION

Isolation and characterization of C-independent (*tin*) mutants

Transcription from the Mu *mom* promoter requires both the activity of the host Dam DNA-methyltransferase and the phage-encoded protein, C. It is known that C is a DNA-binding protein specific for the Mu late promoters (14, 33; Nagaraja et al., in preparation), but its precise role remains to be defined. As with several other prokaryote promoters requiring an accessory transcription factor, the Mu *mom* promoter has relatively poor homology to the consensus *E. coli* promoter sequence. Furthermore, a T_6-run in the spacer may affect transcription due to its intrinsic bending potential (34).

In this paper, we have described the isolation and characterization of two mutations in the *mom*-promoter region which allow C-independent transcription. In both cases, the transcriptional start site (G971) of the C-activated wt promoter was conserved. In *tin7*, a T to G substitution has converted a T_6-run (adjacent to the −10 hexamer) to T_3GT_2, abolishing any intrinsic bending potential of this sequence (34). Additionally, the base substitution created a TG at −15, −14, a situation that has been observed to relieve dependence on accessory transcription factors for a number of other *E. coli* promoters (35, 36), although not in all cases (37). It has been suggested that DNA kinks may be readily introduced at TG/CA base pairs following binding of the *E. coli* CAP protein (38). Thus, it is

possible that RNAP binding to the *tin7* promoter introduces a kink at the TG, and that facilitates formation of new and improved RNAP contacts with residues in the −10 region. If this is so, then it would be worthwhile to determine whether RNAP binding alone induces a kink in those promoters which respond to a TG mutation at −15, −14.

In previous studies we isolated a series of spontaneous *tin* mutations in a *mom-lacZ* multi-copy plasmid (S.Hattman and J.Ives, unpublished results). Many of these contained single base-substitutions scattered throughout the promoter region from −126 to −11. None produced enzyme levels as high as *tin6* or *tin7*, although the specific activity varied over a 20-fold range. These have not yet been examined further, but it is evident that modulation of *mom* expression can be affected by mutations at many sites. Further studies are important to determine the effect of these mutations on RNAP binding at P1 and P2.

Footprinting of RNAP

We have used wt and *tin* derivatives of *com-lacZ* fusion plasmids to investigate the interaction of RNAP with the *mom* promoter region. To this end, DNase I-protection and $KMnO_4$-sensitivity analyses were carried out to probe for the presence of RNAP in the open-complex. The results of these studies are summarized as follows: (i) In the absence of C, RNAP bound to the wt promoter in vivo at a site (designated P2), which overlaps the −35 binding site for C. (ii) In contrast, when C was present in vivo, RNAP bound to the wt promoter at a different site, designated P1, located downstream and partially overlapping P2. (iii) In the absence of C, RNAP bound to the *tin7* promoter predominantly at P1 in vivo and in vitro. With mutant *tin6* (a 63 base-pair deletion removing P2, part of P1 and the C-binding site), RNAP bound to P1 independent of C. (iv) Based on the patterns of protection/sensitivity to DNase I, the boundaries of P1 and P2 are from −56 to +21 and −74 to −24, respectively.

The results of β-galactosidase assays on single-copy *com-lacZ* prophage fusions were consistent with those of the $KMnO_4$-footprinting. The level of β-galactosidase activity for *tin6* was independent of C, and it was as high as the C-activated wt (Table I). On the other hand, the *tin7* promoter was only slightly active in the absence of C, but was transactivated to a level 40 to 50% higher than that of *tin6* or the C-activated wt. These results raise the possibility that the T_6-run plays a role in limiting the level of transcription from the *mom* promoter. Intrinsic curvature at this site in the DNA may occur in a direction that is unfavorable for RNAP binding, and it might magnify the disadvantage of having a 19 nt spacer.

The role of C

In contrast to a previous report (14), we have observed that C binds at other sites in addition to the −35 site; viz. at a site overlapping the −10 hexamer and spacer, as well as two additional sites flanking the −10 and −35 regions (Nagaraja et al., in preparation). We believe that C binds the −10 and −35 sites on the same face of the DNA helix because they are separated by about one helical turn. From the results presented here, we conclude that P1 is the 'functional' RNAP binding site for *mom*-transcription initiation, and that C activation blocks RNAP binding at P2 as well as promoting binding at P1. Considering that the binding sites for the two proteins overlap (probably on different faces of the DNA helix), specific C-RNAP interactions are likely to be involved in the activation process. It seems plausible that C-binding at the −35 site would displace or prevent RNAP binding in P2 (directly by steric hindrance or indirectly by alteration of the DNA conformation). Alternatively, a C-RNAP complex might displace the RNAP bound at P2. It is not known if C binding at the −35 site alone is sufficient to activate transcription, or whether C binding at the −10 site is also required. Since we are only at the beginning of our investigation of *mom* transcriptional activation, further studies are necessary to elucidate the molecular details of these interactions.

What is the nature of the RNAP-P2 complex?

The pattern of $KMnO_4$-hypersites and their rifampicin-dependency in vivo suggest that RNAP bound at P2 (in the wt promoter region) was in an open-complex. Within this region there are several sequences which would probably make relatively poor promoters, thus making the exact placement of P2 difficult. However, in vitro footprinting with DNase I showed that RNAP protected the wt promoter between −74 to −24, whereas the *tin7* promoter was protected from −56 to +21. Accordingly, the best promoter sequence for P2 on the top strand would be TTAAAT ('-35') near −73 and GATTAT ('-10') near −50, with an 18 nt spacer (overlined in Fig. 5). Although RNAP appears to be in an open-complex, in vivo (rightward) transcription from this region was reduced more than 50-fold compared to the C-activated promoter (Fig. 2, lanes 1−3). Therefore, if the RNAP-P2 complex is transcriptionally competent, then the mRNA must be terminated prematurely or rapidly degraded or be in the leftward direction (this would not have been detected because of the nature of the primer employed). In this regard, Margolin et al (13) reported that there is a sequence in the *mom* promoter region that can serve as a C-independent leftward promoter when fused to *lacZ* in the appropriate reading frame; but, detailed in vivo studies of Mu development have failed to show any leftward transcription from this region (39). It is interesting to note that the bottom strand has a TTGAGT ('-35') and a CATAAT ('-10') (underlined in Fig. 5). Although these hexamers have reasonable homology to the consensus, they are separated by a suboptimal spacer length of 13 nt. Clearly, further studies are needed to determine the exact nature of the RNAP-DNA complex at P2.

Why does RNAP bind to the wt *mom* promoter at P2 when C protein is not present? Inasmuch as Mom modification is a 'late' function and one which appears to be cytotoxic if expressed at the wrong time (5), then binding to the P2 site may serve as an additional transcriptional control. That is, P2 may act as a 'sink' for RNAP binding in the absence of C that would minimize chance rightward transcription.

Bacteriophage Mu has evolved a series of intricate mechanisms to regulate gene expression. From the results reported here, there appear to be additional facets of transcriptional regulation; viz. DNA structure and the presence of alternative, functionally distinct binding sites for RNAP influence the availability of the *mom* promoter. In this regard, C probably has dual functions in its role as a transcriptional activator of *mom* transcription; i.e. it prevents binding of RNAP at P2 and activates binding at P1.

ACKNOWLEDGEMENT

This work was supported by a grant from the National Institutes of Health no. GM29227 to S. H.

REFERENCES

1. Reznikoff, W. S., Bertrand, K., Donnelly, C., Krebs, S., Maquat, L. E., Peterson, M., Wray, L., Yin, J. and Yu, X-M. (1987) In W. S. Reznikoff, R. R. Burgess, J. E. Dahlberg, C. A. Gross, M. T. Record, Jr. and M. P. Wickens (eds), RNA polymerase and the regulation of transcription. Elsevier Press. pp105–113.
2. Maeda, S. and Mizuno, T. (1990) *J. Bacteriol.*, **172,** 501–503.
3. Aoyama, T. and Oka, A. (1990) *FEBS Letters*, **263,** 1–4.
4. Greene, J. R., Morrissey, L. M., Foster, L. M. and Geiduschek, E. P. (1986) *J. Biol. Chem.*, **261,** 12820–12828.
5. Kahmann, R. and Hattman, S. (1987) In N. Symonds, A. Toussaint, P. Van de Putte, & M.M Howe (eds), Phage Mu. Cold Spring Harbor Laboratory Press, pp 93–109.
6. Hattman, S. (1982) *Proc. Natl. Acad. Sci. USA* , **79,** 5518–5521.
7. Kahmann, R. (1983) *Cold Spring Harbor Symp. Quant. Biol.*, **47,** 639–646.
8. Hattman, S., Goradia, M., Monaghan, C. and Bukhari, A. I. (1983). *Cold Spring Harbor Symp. Quant. Biol.*, **47,** 647–653.
9. Seiler, A., Blöcker, H., Frank, R. and Kahmann, R. (1986). *EMBO J.*, **5,** 2719–2728.
10. Bölker, M. and Kahmann, R. (1990) *EMBO J.*, **8,** 2403–2410.
11. Hattman, S., Ives, J., Margolin, W. and Howe, M.M. (1985) *Gene*, **39,** 71–76.
12. Heisig, P. and Kahmann, R. (1986) *Gene*, **43,** 59–67.
13. Margolin, W., Rao, G. and Howe, M.M. (1989) *J. Bacteriol.*, **171,** 2003–2018.
14. Bölker, M., Wulczyn, F.G. and Kahmann, R. (1989) *J. Bacteriol.*, **171,** 2019–2027.
15. Sasse-Dwight, S. and Gralla, J. D. (1989). *J. Biol. Chem.*, **264,** 8074–8081.
16. Miller, J. H. (1972). Experiments in Molecular Genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.
17. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
18. Zengel, J. M., Mueckl, D. and Lindahl. L. (1980) *Cell*, **21,** 523–535.
19. Chang, A. Y. C. and Cohen, S. N. (1978) *J. Bacteriol.*, **134,** 1141–1156.
20. Schumann, W., Bade, E.G., Forgie, R.A. and Howe, M.M. (1980) *Virology*, **104,** 418–425.
21. Birnboim, H.C. and Doly, J. (1979) *Nucleic Acids Res.*, **7,** 1513–1523.
22. Minton, N.P. (1984) *Gene*, **31,** 269–273.
23. Tiedeman, A. T. and Smith, J. M. (1988) *Nucleic Acids Res.*, **16,** 3587–3587.
24. Simons, R. , Houman, W. F. and Kleckner, N. (1987). *Gene*, **53,** 85–96.
25. Lamond, A. I. and Travers, A. A. (1985) *Cell* , **40,** 319–326.
26. Knaus, R. and Bujard, H. (1988). *EMBO J.*, **7,** 2919–2923.
27. Sasse-Dwight, S. and J. D. Gralla. (1988) *J. Mol. Biol.*, **202,** 107–119.
28. Gralla, J. D. (1985) *Proc. Natl. Acad.Sci.USA*, **82,** 3078–3081.
29. Hattman, S. and Ives, J. (1984) *Gene*, **29,** 185–198.
30. Plasterk, R. H. A., Vollering, M., Brinkman, A. and Van de Putte, P. (1983). *Cell*, **36,** 189–196.
31. Sasse-Dwight, S. and J. D. Gralla. (1988) *Proc. Natl. Acad.Sci.USA*, **85,** 8934–8938.
32. Galas, D. J. and Schmitz, A. (1978) *Nucleic Acids Res.*, **5,** 3157–3170.
33. Nagaraja, V., Hecht, G. & Hattman, S. (1988) *Biochem. Pharmacol.*, **37,** 1809–1810.
34. Koo, H-S., Wu, H-M. and Crothers, D. M. (1986) *Nature*, **320,** 501–506.
35. Keilty, S. and Rosenberg, M. (1987) *J. Biol. Chem.*, **262,** 6389–6395.
36. Ponnambalam, S. and Busby, S. (1988) *Mol. Microbiol.*, **2,** 165–172.
37. Bracco, L., Kotlarz, D., Kolb, A., Diekmann, S. and Buc, H. (1989) *EMBO J*, **8,** 4289–4296.
38. Schultz, S. S., Shields, G. C. and Steitz, T. A. (1991) *Science*, **253,** 1001–1007.
39. Marrs, C F. and Howe, M. M. (1990) *Virology*, **174,** 192–203.