JOURNAL OF BACTERIOLOGY, Mar. 1989, p. 1623–1630
0021-9193/89/031623-08$02.00/0

Vol. 171, No. 3

Cyclic-AMP-Dependent Switch in Initiation of Transcription from the Two Promoters of the *Escherichia coli gal* Operon: Identification and Assay of 5′-Triphosphate Ends of mRNA by GTP:RNA Guanyltransferase

M. IRANI,† R. MUSSO,[1] AND S. ADHYA[2]*

Laboratory of Molecular Biology, National Cancer Institute, Bethesda, Maryland 20892,[2] and Department of Biology, University of South Carolina, Columbia, South Carolina 29208[1]

Received 18 August 1988/Accepted 19 December 1988

We have studied the initiation of transcription of the *gal* operon in *Escherichia coli* (i) by analyzing the 5′-triphosphate ends and (ii) by measuring the level of promoter-proximal *gal* mRNA made in vivo. The 5′ termini were identified and quantified by capping with GTP:mRNA guanyltransferase, and the mRNA levels were determined by hybridization of pulse-labeled [^{32}P]RNA with a specific DNA probe. Our results conclusively demonstrate the in vivo activities of two promoters, P1 and P2, with separate initiation sites (S1 and S2) as suggested before from in vitro and in vivo experiments (S. Adhya and W. Miller, Nature [London] 279:492–494, 1979; R. E. Musso, R. DiLauro, S. Adhya, and B. de Crombrugghe, Cell 12:847–854, 1977). We have also studied the effect of cyclic AMP (cAMP) on in vivo *gal* transcription and found that whereas total *gal* transcription remains largely unchanged, the relative proportions of the S1 and S2 mRNAs are influenced by the level of cAMP in the cell. In strains devoid of cAMP (*cya*), transcription initiates equally at S1 and S2; in cAMP-proficient cells (*cya*$^+$), the S1 initiation increases twofold with a concomitant decrease in S2 initiation. Addition of a saturating amount of exogenous cAMP to *cya* mutant cells results in a relatively larger switch from S2 to S1. Our results clearly show that while cAMP is an inhibitor of S2, it is not an absolute requirement for transcription initiation at S1, but only acts to increase low-level transcription from the P1 promoter. Using these approaches, we have also studied *gal* promoter mutants (*P211*, *P18*, and *P35*) which show altered behavior in transcription initiations and in response to cAMP. On the basis of these results, we have discussed models by which transcription initiates at the two overlapping *gal* promoters (P1 and P2) and discussed how cAMP level modulates the switch between them.

From genetic analysis and in vitro transcription studies of the *gal* operon of *Escherichia coli*, we have postulated that the operon is under the control of two partially overlapping promoters, P1 and P2, having separate start sites, S1 and S2, which are 5 base pairs (bp) apart (2, 17). The structure and DNA sequence of the control region of the operon are shown in Fig. 1. S1 is designated as +1. Cyclic AMP (cAMP) and its receptor protein (CRP) have opposite effects on the *gal* promoters. The cAMP · CRP complex acts as a positive control element for P1 and as a negative factor for P2. Genetic analysis and in vitro chemical protection studies have localized the initial binding site of cAMP · CRP around position −40 on the DNA, with the G residue at position −35 probably making contact with the protein (5, 6, 27; A. Majumdar and S. Adhya, manuscript in preparation). It is believed that cAMP · CRP performs its dual role by binding to this site, although it has been suggested more recently that a second molecule of CRP can bind in a sequence-nonspecific way at the −60 region when RNA polymerase occupies P1 (8, 20, 21, 26).

By analyzing the *gal* mRNA by S1 nuclease protection studies, we have previously shown that the same two transcription start sites, S1 and S2, are likely to operate in vivo (4). However, this sizing assay does not distinguish between intact and processed RNA molecules and thus should not be taken as a rigorous criterion for determining the original 5′ terminus of an mRNA. In this paper, we conclusively demonstrate the existence of two start sites of *gal* operon transcription by capping the 5′-triphosphate ends and identifying the capped *gal* mRNA. Using the wild type and several mutant promoters, we have also studied the influence of cAMP · CRP in turning P1 on and P2 off as initiation of *gal* transcription switches between the two promoters.

MATERIALS AND METHODS

Materials. The bacterial strains used in this study are described in Table 1. The λ *gal*8 phage used is described by Feiss et al. (10). Plasmids pMI1 and pMI4 are described below. M56 minimal medium (2), containing 0.2% D-fructose as a carbon source and supplemented with 20 μg each of L-histidine, L-isoleucine, and L-valine per ml and 1 μg of vitamin B_1 per ml, was used. Exogenous cAMP, when added, was at a concentration of 10 mM. SSC (1×) contains 0.15 M NaCl and 0.015 M sodium citrate.

GTP:mRNA guanyltransferase (capping enzyme), purified from vaccinia virus (28), and poly(A) were gifts of S. Venkatesan. RNase T_1 was from Calbiochem, and RNase A and DNase I (RNase free) were from Worthington.

[α-^{32}P]GTP (300 to 400 Ci/mmol) was obtained from Amersham Corp., and [5-^{3}H]uridine (>20 Ci/mmol) was obtained from Schwarz BioResearch. Nitrocellulose membrane filters were supplied by Schleicher & Schuell.

Methods. (i) Transfer of *galP* mutants from plasmid to chromosome. The *gal* plasmids used in this study have been

* Corresponding author.
† Present address: ZymoGenetics, Seattle, WA 98105.

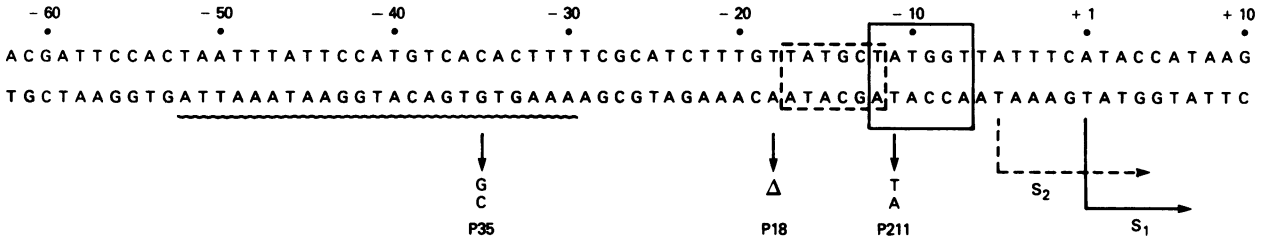

FIG. 1. Structure and sequence of the *gal* promoter region. The transcription start sites S1 (for the cAMP · CRP-dependent transcript) and S2 (for the cAMP · CRP-independent transcript) are designated +1 and −5, respectively; the corresponding −10 and −35 regions are boxed, as adopted from Musso et al. (17). The mutational changes associated with the three mutants studied, *galP211* (17), *galP18*, and *galP35* (5, 6), are also indicated. The wavy underline shows the sequence for cAMP · CRP binding.

described earlier (6) and were supplied by S. Busby. The *gal* mutations were transferred first to λ *gal*8 transducing phage and then to the bacterial chromosome as described in an earlier communication (14).

(ii) Preparation of plasmid DNA for hybridization. DNA for each plasmid was prepared by lysing cells with Triton (30) and purified by banding twice in CsCl-ethidium bromide gradients. The Sepharose B chromatography step was omitted. For the quantitative determination of promoter-proximal labeled *galE* mRNA, the linearized plasmid DNAs used as probes were further purified by isolating the linearized DNA bands from 0.8% agarose gels by the perchlorate method (7).

For purification of *gal* mRNA for capping, the *l*-strand of λ *gal*8 phage DNA, prepared by the method of Szybalski et al. (25), was supplied by D. Court.

For DNA hybridization, 7-mm filter disks were punched out of large filters. The amount of DNA loaded onto a large filter was adjusted to yield the desired amount of DNA in each punched-out filter. Alkaline-denatured, linearized plasmid DNA was diluted to concentrations between 1 and 5 μg/ml with 10 ml of 6× SSC, neutralized, and passed dropwise through previously soaked and washed filters. The filters were again washed with 6× SSC solution under high suction, dried at room temperature, punched, and baked at 72°C for 2 h under vacuum. The filters were used within 1 week after baking. In loading the *l*-strand DNA of phage λ *gal*8, 25-mm filters were used and the alkaline denaturation step was omitted.

(iii) Isolation of *gal* mRNA for capping. An overnight culture of cells grown in supplemented M56 medium containing fructose was diluted 345-fold into 70 ml of fresh medium, additionally containing 0.2% D-galactose as an inducer of the *gal* operon, and was grown to an optical density of 0.3 at 590 nm at 37°C. The cells were poured over crushed ice and sodium azide (0.1 M), collected, and lysed in 8 ml of lysis buffer (0.01 M Tris hydrochloride, pH 7.5, 0.01 M NaN_3, 0.001 M EDTA, and 200 μg of lysozyme per ml) by two cycles of freezing and thawing. After treatment with 20 μg of DNase (RNase free) per ml in the presence of 10 mM $MgCl_2$ on ice for 20 min, acetic acid and sodium dodecyl sulfate (SDS) were added to 4 mM and 0.5%, respectively, and the resulting clear lysate was extracted twice with an equal volume of a 1:1 mixture of phenol and chloroform. The latter mixture was previously equilibrated with 0.02 M sodium acetate, 0.02 M potassium acetate, and 0.01 M magnesium acetate at pH 5.2. The RNA was purified by two cycles of precipitation with ethanol, washed, dried, and dissolved in water.

Samples of pMI4 DNA (10 μg; see Results), immobilized on 7-mm filter disks, were used to purify the *gal* mRNA by hybridization. Samples of RNA (1.5 mg) were used to hybridize to DNA filters in 600 μl of hybridization buffer consisting of 20 mM PIPES buffer [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)], 400 mM NaCl, 50% formamide (deionized), and 0.1% SDS at pH 6.5 at 37°C for 16 h. The reaction mixture was then chilled, and the filters were washed at room temperature six times with 10 ml of 2× SSC, followed by two incubations of 30 min each at 37°C in the hybridization buffer without SDS. For elution, the washed filters were incubated at 67°C for 6 min with 90% formamide in 20 mM PIPES buffer at pH 6.5, containing 30 μg of alkaline phosphatase-treated tRNA. The incubation was repeated with fresh solution without tRNA, and the RNA was purified from the combined supernatants by three cycles of ethanol precipitation. In later experiments, *l*-strand DNA of λ *gal*8 phage immobilized on 25-mm filters was used as the *gal*-specific probe.

(iv) End labeling of *gal* mRNA. Capping of RNA was performed essentially as described by Venkatesan and Moss (29). The RNA recovered after hybridization was incubated in 50-μl reactions containing 50 mM Tris hydrochloride (pH 7.5), 4 mM dithiothreitol, 4 mM $MgCl_2$, 50 μM *S*-adenosylhomocysteine, 300 μCi of [α-^{32}P]GTP, and 6 μl of RNA guanyltransferase. After 30 min at 37°C the reactions were terminated by the addition of EDTA to 10 mM and SDS to 0.2%. The RNA was extracted with phenol-chloroform (1:1), ethanol precipitated, and dissolved in 50 μl of water. It was then passed through a column (0.7 by 10 cm) of G-50 Sephadex to remove the residual GTP and was ethanol precipitated twice. This capped RNA was rehybridized to pMI4 DNA filters in 200 μl of hybridization buffers as

TABLE 1. *E. coli* K-12 strains used

Strain	Genotype	Source or reference
SA500	F⁻ *thi his rpsL relA*	(3)
SA2346	SA500, *ilv-2*::Tn*10* Δ*cya*[a]	This study
SA2231	SA500, *ilv-2*::Tn*10 cya-1039*	This study
MI104	SA500, Δ(*gal-chlA*)*1421 lacI lacPUV5*	This study
MI601	MI104, *galP*⁺	This study
MI701	MI104, *galP35*	This study
MI703	MI104, *galP18*	This study
MI705	MI104, *galP211*	This study
MI602	MI601, *galP*⁺ *ilv*::Tn*10* Δ*cya*	This study
MI604	MI601, *galP*⁺ *ilv*::Tn*10 cya-1039*	This study
MI702	MI701, *galP35 ilv*::Tn*10* Δ*cya*	This study
MI704	MI703, *galP18 ilv*::Tn*10* Δ*cya*	This study
MI706	MI705, *galP211 ilv*::Tn*10* Δ*cya*	This study

[a] Δ*cya* allele originated from strain CA8306 obtained from J. Beckwith.

described above. For elution, the filters were washed with 2× SSC six times at room temperature, eluted either with 90% formamide, as described above, or by heating with water in the presence of 50 μg of tRNA at 90°C, and finally recovered by ethanol precipitation.

(v) Pulse-labeling of RNA with [^{3}H]uridine. Cells grown to an A_{590} of 0.3 as described above were exposed to 50 μCi of [^{3}H]uridine per ml for the desired length of time and then arrested by pouring over a mixture of ice and sodium azide (0.1 M). The RNA was extracted using diethyl pyrocarbonate essentially as described by Summers (24). It was purified by two ethanol precipitations. The specific activities of the RNA samples ranged from 20,000 to 115,000 cpm/μg of RNA for durations of the pulse ranging between 15 and 90 s.

(vi) Hybridization assay of pulse-labeled *galE* message. The hybridizations were performed in 200 μl of 2× SSC at 67°C for 20 h in the presence of 0.1% SDS and 200 μg of yeast tRNA. Filters (7 mm) containing 0.5 μg of highly purified plasmid DNA were used. The reactions were terminated by chilling on ice. The filters were then washed six times with 3 ml each of 2× SSC and incubated with 20 μg of RNase (heated for 10 min at 90°C to inactivate DNase) in 1 ml of 2× SSC at room temperature for 1 h. They were rewashed eight times with 2× SSC, dried under a heating lamp for 10 min, and counted in Econofluor in a Beckman scintillation counter. For estimating the background hybridization, a similarly prepared filter containing 0.5 μg of purified pBR322 DNA was incubated in each of the hybridization tubes. The *gal*-specific counts were taken to be the difference between the counts retained by the *gal*-DNA filters and those retained on the pBR322 filters. The counts on the control filters were approximately 10% of the experimental counts. In kinetics experiments (see Table 2), a fixed amount of [^{3}H]RNA (20 μg) was used as input RNA for each of the time points.

(vii) Computer analysis of promoter sequences. The DNA sequences of the *gal* wild type and of promoter mutants were compared with a consensus promoter sequence by using the computer program developed by Mulligan et al. (16). Only those promoterlike sequences resulting in RNA start sites at +1 (i.e., *galP1*) or −5 (i.e, *galP2*) are reported; several pseudo-promoter sequences with homology scores of 30 to 45% were detected by the computer but not by experimental analysis of RNA transcripts.

RESULTS

Mapping of the 5′ end of in vivo *gal* mRNA. The principle of the mapping technique is outlined in Fig. 2. Total RNA isolated from exponentially growing *E. coli* cells was specifically end labeled with [α-^{32}P]GTP at the 5′-triphosphate termini. The reaction catalyzed by the enzyme GTP:mRNA guanyltransferase involves an initial cleavage of the terminal 5′-tri- or diphosphate of the RNA molecules, followed by the transfer of the [^{32}P]GMP residue from GTP to the 5′ end of RNA, forming a cap structure (28). In the second step, the capped RNA is digested with RNase T_1, generating labeled 5′-terminal T_1 oligonucleotides whose size would be characteristic of the position of transcription initiation with respect to the position of the first GMP residue in the mRNA.

We deduced the in vivo initiation sites of *gal* transcription on the basis of the size analysis of the terminal T_1 oligonucleotides from *gal* mRNA. *gal* transcription in vitro initiates at +1 (Fig. 1) in the presence of cAMP and CRP and at −5 in their absence (17). If the same two start sites are functional in vivo, as suggested by protection experiments with S1 nuclease (4), the S1 transcript would generate a T_1 decamer

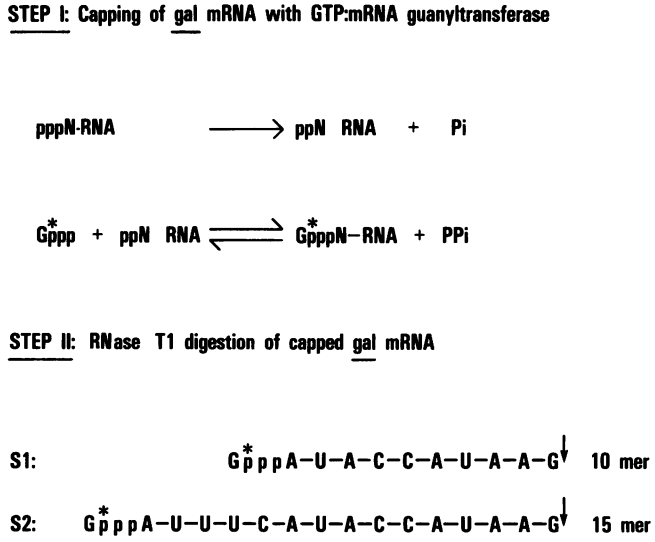

FIG. 2. Scheme for mapping the 5′-triphosphate ends of *gal* mRNA. The asterisks indicate the phosphate labeled with ^{32}P. The downward arrows are the sites of digestion by RNase T_1.

with the cap structure on the 5′ end, and S2 transcript would produce a capped T_1 oligonucleotide 15 bases long. The two oligonucleotides can easily be resolved by gel electrophoresis, and their size can be determined by comparing their mobilities with those of capped poly(A) ladder. Figure 3 shows the results of such an analysis performed on RNA samples from wild-type and adenylate cyclase-defective

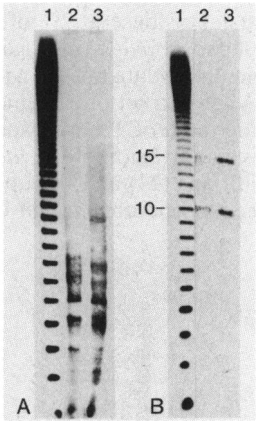

FIG. 3. 5′-End analysis of *gal* mRNA. (A) *gal* mRNA, cap labeled as described in Materials and Methods, was digested with RNase T_1 in 10 μl of 20 mM Tris hydrochloride (pH 7.4)–2 mM EDTA for 45 min at 37°C. The samples were next dried in vacuo, dissolved in 4 μl of formamide-dye solution (15), and fractionated by electrophoresis on 25% polyacrylamide–8 M urea gels by the procedure of Maxam and Gilbert (15). Results with RNA from MI601 (*cya*$^+$) (lane 2) and with RNA from MI604 (*cya-1039*) (lane 3) are shown. The calibration standard of capped poly(A) degraded by heating with formamide is displayed in lane 1. The 10- and 15-mer capped poly(A) residues are marked. Salt concentration and total amount of RNA were adjusted to match the experimental samples. (B) The capped RNA was purified by one more round of hybridization to *gal* DNA (Materials and Methods) before being digested with RNase T_1. The conditions of RNase T_1 treatment and electrophoresis and the designation are the same as for panel A. Lane 1, Capped poly(A) residues; lane 2, *cya*$^+$ RNA; lane 3, *cya* RNA. The total counts loaded in each lane vary somewhat.

mutant (*cya*) cells. *cya* cells do not make cAMP and are expected to increase the level of S2 RNA while decreasing S1 RNA (see next section). In initial experiments, the *gal* mRNA was isolated following a hybridization of total cellular RNA to *gal* DNA as described in Materials and Methods. The results (Fig. 3A) show that the capped material after T_1 digestion produced oligonucleotides ranging from the smallest to about 12 bases long, indicating that the *gal* mRNA from both *cya*$^+$ and *cya* cells was heavily contaminated with other mRNA species. However, complete purification was achieved by a second cycle hybridization to *gal* DNA, after capping but before the *gal* mRNA was subjected to T_1 digestion. Less than 1% of the counts came through a second hybridization; these appeared, after treatment with RNase T_1, as two bands at positions 10 and 15 (Fig. 3B, lanes 2 and 3). This clearly shows that the mRNA species which hybridizes specifically to the cloned *gal* DNA initiates at two distinct positions which correspond precisely to the S1 and S2 positions defined in vitro (see Fig. 1). The plasmid pMI4 used in the *gal*-specific hybridization carries the first 374 base pairs of *gal* mRNA and a 187-bp region upstream of the S1 start site. Inspection of the DNA sequence of this region does not reveal any other potential promoterlike sequences, which could account for the existence of the 10- and 15-mer RNase T_1-generated capped oligonucleotides. The relative amounts of S1 and S2 in wild-type and *cya* cells are discussed next.

Effect of cAMP on *gal* initiation in vivo. The relative proportions of the two types of initiations in *gal* in wild-type and various mutant cells (see Fig. 3 to 5) were studied by quantifying the two capped oligonucleotides (Table 2). We have used two *cya* mutations, a point mutation (*cya-1039*) and a deletion (Δ*cya*), for this purpose. Both mutants, unlike the wild-type strain (*cya*$^+$), are pleiotropic sugar defective and synthesize β-galactosidase at 2% of the wild-type level in the presence of isopropyl-β-D-thiogalactopyranoside in the experimental medium (data not shown; see reference 13). For the *gal* mRNA assay a set of isogenic strains, harboring additionally *lacI* and the *lacUV5* mutations, were used. This ensured constitutive expression of the *lacY* gene under both *cya*$^+$ and *cya* conditions (23) and the uptake of D-galactose for induction of the *gal* operon through LacY permease (1).

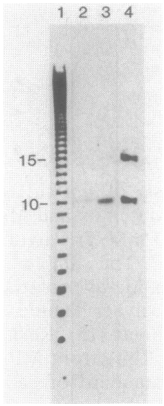

FIG. 4. Effect of cAMP on *gal* initiation in vivo. The experimental conditions were the same as described in Fig. 3B except that the first hybridization was to the *l*-strand of λ *gal*8. The following RNAs were analyzed: lane 2, from MI601 (*cya*$^+$); lane 3, from MI602 (Δ*cya*) grown in the presence of 10 mM cAMP; lane 4, from MI602 (Δ*cya*). The numbers 10 and 15 indicate the positions of the capped oligonucleotides 10 and 15 bases long, respectively. Lane 1 is again the poly(A) standard.

TABLE 2. Modulation of P1 and P2 by cAMP

Relevant features	Relative amts[a] of:		% *galE* message[b]	Relative amts of *galE* message from[c]:	
	S1	S2		P1	P2
1. *cya*$^+$ *galP*$^+$	2	1	100	66	33
2. *cya-1039 galP*$^+$	1	1	ND[d]		
3. Δ*cya galP*$^+$	1	1	88	44	44
4. Δ*cya galP*$^+$ (+cAMP)	4	1	100	80	20
5. *cya*$^+$ *galP35*	1	3	60	15	45
6. Δ*cya galP35*	(1)	(3)	59	(14)	(45)
7. *cya*$^+$ *galP18* (P2$^-$)	+	−	25	25	0
8. Δ*cya galP18* (P2$^-$)	+	−	46	46	0
9. *cya*$^+$ *galP211* (P1$^-$)	−	+	31	0	31
10. Δ*cya galP211* (P1$^-$)	−	(+)	40	(0)	(40)

[a] The relative amounts of *gal* RNA starting at S1 and S2 were derived by densitometry scanning of autoradiograms shown in Fig. 3B, 4, and 5. The relative proportions were determined by weighing the area of paper corresponding to the two peaks. Numbers in parentheses are very approximate due to low amounts of radioactivity in these analyses. When only one RNA was detectable, the results are indicated by + and −.

[b] The fraction of total RNA synthesis corresponding to *galE* mRNA is expressed as a percentage of that determined with control strain MI601 (taken as 100%). Results are an average of at least two determinations. For strains MI601, MI602, and MI701, RNA was prepared from culture samples taken at various times after a 90-s pulse-labeling with [^{3}H]uridine. The *gal*-specific RNA was determined by hybridization of the first 400 nucleotides of *galE* (see Materials and Methods) and plotted as a function of total trichloroacetic acid-precipitable ^{3}H counts. The slope of the resulting linear graph measured the *galE* transcription as a fraction of total incorporation of [^{3}H]uridine into extractable RNA. For the remaining strains, cultures were pulse-labeled for 40 s, and *galE* mRNA was assayed in duplicate samples and normalized to total incorporation.

[c] The relative level of *gal* transcription from P1 and P2 is calculated from the amount of total *galE* mRNA synthesis determined as described in footnote *b* above and the relative utilization of S1 and S2 start sites determined as described in footnote *a*. Owing to the limited precision in the latter, minor differences in the final calculated levels for P1 and P2 transcription are not regarded as significant.

[d] ND, Not determined.

As we have discussed before, it was apparent (Fig. 3B) that *gal* transcription initiates at S1 and S2 in both the wild type and the *cya-1039* mutant. In the *cya-1039* mutant both the start sites were utilized with equal efficiency (Table 2, line 2). This changed to a ratio of 2:1 in favor of the S1 transcript in cAMP-proficient wild-type cells (line 1).

Figure 4 (lane 4) shows the amounts of S1 and S2 in the Δ*cya* strain. These cells behaved identically to the point mutant, giving an equal proportion of the capped oligonucleotides (Table 2, line 3). Thus *gal* transcription from P1 cannot be explained by residual cAMP due to leakiness of the *cya-1039* point mutation. Lane 2 of Fig. 4 displays the 2:1 distribution of S1 to S2 for the wild-type strain. If the higher ratio of S1 to S2 in the wild-type strain is a consequence of higher cAMP levels in these cells, as predicted before (2, 17), the addition of the cyclic nucleotide to the *cya* cultures should elicit the same response. In the presence of exogenous cAMP, the Δ*cya* strain changed the ratio of S1 to S2 from 1:1 (Fig. 4, lane 4) to 4:1 (Fig. 4, lane 3) (Table 2). The higher (4:1) ratio in this case compared to that of 2:1 in the *cya*$^+$ strain was due to the subsaturating levels of cAMP present in wild-type cells growing under the conditions used (data not shown).

Thus, from the results of Fig. 3B and 4, it is apparent that *gal* transcription initiates at S1 and S2 in both the presence and absence of cAMP. The ratio of S1 to S2 is about 1:1 in the absence of cAMP and changes to about 4:1 when the cyclic nucleotide is in excess.

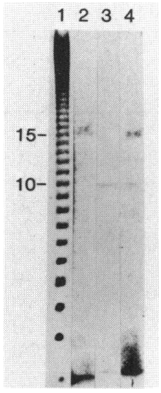

FIG. 5. Analysis of *gal* initiation in *galP* mutants. The experimental conditions were the same as described for Fig. 3B. Analysis of *gal* transcription in (lane 2) MI705 (*galP211 cya*$^+$), (lane 3) MI704 (*galP18* Δ*cya*), and (lane 4) MI701 (*galP35 cya*$^+$). Lane 1 is the calibration standard, with the numbers 10 and 15 marking the positions of capped poly(A) molecules 10 and 15 residues long, respectively.

5′-End analysis of *gal* mRNA in various *galP* mutants. From the preceding section it is clear that the transcription of the *gal* operon is initiated at two start sites, which are modulated in opposite directions by the level of cAMP in the cell. Several promoter variants of the *gal* operon have been characterized for their pattern of transcription and for their ability to bind cAMP · CRP in vitro (3, 5, 6, 17, 22). We have investigated the in vivo initiations of some of these mutants. The three *galP* mutations (*galP211*, *galP18*, and *galP35*) studied are shown in Fig. 1. Of these, *galP18* and *galP35* were isolated in recombinant plasmids and were transferred to the chromosome as explained in Materials and Methods.

(i) *galP211*. The *galP211* mutant carries an A · T-to-T · A transversion at position −11, which inactivates the P1 promoter but has only a marginal effect on P2 (2, 17, 22). In contrast to the wild-type promoter, the *galP211* mutant yields capped S2 RNA but not S1 RNA in both *cya*$^+$ (Fig. 5, lane 2) and *cya* (not shown) strains. This confirms the prediction that the *galP211* mutation affects the P1 promoter, which is stimulated by cAMP, and not P2, which is inhibited by the cyclic nucleotide.

(ii) *galP18* and *galP35*. *galP18* is a single-base-pair (T · A) deletion at position −18. This mutant does not make the S2 transcript in vitro, but shows S1 transcript in the absence but not in the presence of cAMP (6). In vivo detection of capped *gal* RNA in this mutant shows only S1 RNA in both Δ*cya* (Fig. 5, lane 3) and *cya*$^+$ (data not shown) cells. Opposite results were obtained with the *galP35* mutant in which the cAMP · CRP-binding site had been mutated, resulting in abolition of P1 transcription in vitro (5, 6). In vivo, the *galP35* mutation initiates *gal* transcription largely at S2, both in *cya*$^+$ (Fig. 5, lane 4) and in *cya* (not shown) cells. If the cAMP · CRP complex fails to bind to the *galP35* DNA in vivo, then the normal 2:1 ratio of S1 to S2 for the wild-type promoter should be skewed in the direction of S2 in a *cya*$^+$ background. The observed distribution of S1 to S2 was 1:3 (Table 2), which confirms this prediction.

Rate of synthesis of *galE* mRNA. cAMP changes the relative proportions of the two starts of *gal* transcription, as measured above by quantifying the 5′-triphosphate ends. We have also studied the rate of total *gal* mRNA elongation (from both start sites) by estimating the amount of pulse-labeled RNA which is hybridizable to *gal* DNA and normalizing that to the net amount of radioactivity in total extractable RNA, as described in Materials and Methods. For a *gal*-specific hybridization probe, we used plasmid pMI4 DNA, which contains the promoter-proximal 400 bp downstream from the +1 start (S1) and the 187 bp upstream of it (Table 2, column 3). The rate of *galE* message synthesis in the *galP*$^+$ *cya*$^+$ strain (Table 2, line 1) is compared with that in *galP*$^+$ Δ*cya* cells growing in the presence (line 3) and absence (line 4) of cAMP. The results show that the rate of total *galE* transcription was not very different (88 versus 100%) in the *galP*$^+$ strains under conditions which altered the ratio of S1 to S2 from 1:1 to 4:1. The cAMP-deficient cells showed only a 12% decrease. In contrast, in the three *galP* mutants examined, there was a more significant reduction of total *galE* RNA transcribed (as low as 25%). These results are discussed later.

To verify that the measured RNA was purely *gal* mRNA, we hybridized the labeled RNA to a shorter *gal* DNA probe, pMI1, which carries the same upstream region but only 45 bp downstream of the +1 site. The radioactive RNA hybridized to the shorter plasmid was 1/10 of that observed with pMI4, indicating that the hybridizable RNA is mostly *gal* mRNA and the upstream DNA present on the plasmids is not significantly transcribed under the conditions of *gal* induction used.

The last column of Table 2 displays the distribution of the total *galE* transcription between the two initiations according to the ratios of S1 and S2 derived from the analysis of the capped T_1 oligonucleotides. This provides a means for comparing the changes in the actual elongation of RNA from each promoter under different sets of conditions. It is now clear that cAMP and CRP, in vivo as in vitro, act both to stimulate transcription from P1 and to inhibit that from P2 (Table 2, line 3 versus line 4). A fourfold increase in the ratio of S1 to S2 translates to a twofold activation of P1 (44 versus 80%) and a twofold inhibition of P2 (44 versus 20%).

With the *galP35* mutant which does not bind cAMP · CRP, the P2 transcription in the Δ*cya* background (45%) was the same as that observed with *galP*$^+$ Δ*cya* (44%) (Table 2, line 5 versus line 3) cells. The activity of P1 (15%), however, was significantly lower (about threefold) than that seen in *galP*$^+$ in Δ*cya* cells (44%), suggesting that the mutation may also have affected the cAMP-independent P1 activity.

The *galP18* mutation did not make S2-initiated RNA in either *cya*$^+$ or *cya* strains (Table 2, lines 7 and 8). The amount of S1 transcript was low (25%) in the presence of cAMP but was considerably higher (46%) in the absence of the cyclic nucleotide. The amount of S1 RNA synthesis and its inhibition by cAMP in *galP18* paralleled that of S2 RNA in the wild-type promoter, as seen in vitro.

When the *galP211* mutation was characterized for relative promoter strength, we found that the amount of S2-initiated transcript was very comparable to that of the wild-type promoter in the *cya*$^+$ background (31 versus 33%) as well as in the Δ*cya* background (40 versus 44%). There was no detectable S1 RNA under either condition. Thus, cAMP does not cause P1 activation but displays P2 inhibition.

DISCUSSION

The cAMP · CRP complex binds to the promoters of many operons, e.g., *lac*, *srl*, etc., and dramatically stimulates their rate of transcription. In the absence of cAMP, the residual transcription of these operons is very low. Consequently, a *cya* strain is phenotypically Lac$^-$, Srl$^-$, etc. Under appro-

TABLE 3. Homology scores of wild-type and mutant *gal* promoters

Promoter	Mutation	−35 Region	Spacing −35/−10 (bp)	−10 Region	Start site	Homology score[a]
P1	Wild type	GTCACA	21	TATGGT	A(+1)	40.8%
P1	Wild type	CACTTT[b]	17	TATGGT	A(+1)	ND
P2	Wild type	ATGTCA	18	TATGCT	A(−5)	59.8%
P1	*galP211*	GTCACA	21	TTTGGT	A(+1)	29.6%
P2	*galP211*	ATGTCA	18	TATGCT	A(−5)	59.2%
P1	*galP35*	GTCAGA	21	TATGGT	A(+1)	35.5%
P2	*galP35*	ATGTCA	18	TATGCT	A(−5)	58.6%
P1	*galP18*	GTCACA	20	TATGGT	A(+1)	40.8%
P1	*galP18*	ATGTCA[c]	22	TATGGT	A(+1)	ND
P2	*galP18*	ATGTCA	17	TATGCT	A(−5)	60.4%

[a] Homology scores were determined by the program of Mulligan et al. (16) and reflect sequences besides the −35 (ideally TTGACA), −10 (ideally TATAAT), and spacing (ideally 17 bp) determinants shown above. ND, Not detected.

[b] An alternative −35 assignment which has been previously suggested for promoter P1 but which is not detected by the Mulligan program (homology, <30%) due to the poor match to the consensus −35 sequence.

[c] An alternative −35 assignment suggested by Busby et al. (5) for P1 of *galP18* but which is not recognized by the Mulligan program (homology score, <30%) due to the excessively long −35/−10 spacing.

priate conditions, in vitro *gal* transcription can also be shown to be dependent on the cAMP · CRP complex (18), and the specific binding of cAMP · CRP to the *gal* promoter region has been demonstrated by DNase protection experiments (27; Majumdar and Adhya, in preparation; see Fig. 1). However, a *cya* strain is phenotypically Gal$^+$ (2). This apparent discrepancy has been explained by postulating the presence of two promoters in *gal*, P1 and P2. The former is dependent upon cAMP, and the latter is inhibited by it. Thus, *gal* is transcribed from P1 in a *cya*$^+$ strain and from P2 in a *cya* mutant. With a purified transcription system, we have in fact discovered two start sites of transcription in *gal*, S1 and S2 (17). The initiation at S1 is stimulated by cAMP · CRP, and that at S2 is inhibited. In qualitative agreement with these observations, we have shown here that in vivo *gal* transcription starts with the same two 5′-triphosphate ends as in vitro. Using *cya* strains we have shown that cAMP · CRP enhances initiation at S1 and inhibits starts at S2. Comparison of the start sites for wild type and several mutant *gal* promoters in both *cya*$^+$ and *cya* cells demonstrated that in vivo S1 and S2 correspond, respectively, to P1 and P2. The results also authenticate the use of capping techniques in identifying true transcription start sites.

Our results also demonstrate that both S1 and S2 start in the absence of cAMP with nearly equal efficiency. Such cAMP-independent transcription has also been seen in vitro (12, 17), although not under some conditions (18). According to the consensus sequences developed for promoters (11), P1 has a good −10 spacing but either a poor −35 or an excessive −10 to −35 spacing, and P2 has a fair −35 but a poor −10 spacing (Table 3). Equally efficient transcription from both P1 and P2 in the absence of cAMP can be explained in two ways: (i) RNA polymerase forms a closed complex at the good −35 of P2 and, during subsequent isomerization to open complex, switches over to P1 about half of the time, or (ii) RNA polymerase forms closed complexes and isomerizes to open complexes at P1 and P2 independently. In the latter event, the net efficiency is similar for the two promoters. Using a computer program to compare DNA sequences, Mulligan et al. (16) have calculated homology scores of 40.8 and 59.8 for P1 and P2, respectively, which reflect their promoter strengths (Table 3). Recent kinetic studies of open complex formation in *gal* support model (ii) above (12). These in vitro studies also indicate that interconversions between open complexes are slow processes and probably not significant.

In vivo, we observed that the presence of cAMP shifted the relative proportion of S1 to S2. With a *cya* mutant, that ratio of S1 to S2 changed from 1:1 (no cAMP) to 4:1 (excess cAMP). A ratio of 2:1 was seen in wild-type cells, in which the cAMP level is less than maximal when growing in fructose plus galactose minimal medium. The results of transcription initiation and extension in *gal* show that P1 is different from the cAMP-dependent *lac* promoter. The residual *lac* transcription is very low and cAMP stimulates it 50-fold, whereas there is considerable *gal* transcription from P1 without the cyclic nucleotide which stimulates it only twofold. Concurrently, cAMP reduces the *gal* transcription from P2 also by twofold.

cAMP · CRP binds to *gal* DNA at the −30 to −52 region, which contains the 5′-AANTGTGANNT consensus sequence developed by Ebright et al. (9). It overlaps with the −35 region for P2 as well as with that for P1. Thus, it is very likely that cAMP · CRP reduces P2 activity by competing with RNA polymerase for closed complex formation at P2. The twofold inhibition of P2 by cAMP in vivo may reflect the relative binding affinities of cAMP · CRP and RNA polymerase competing for the same site and their concentrations in the cell. In this model, only a portion of the *gal* promoters in the population will be occupied by the cyclic nucleotide complex in wild-type cells. The bound cAMP · CRP would then stimulate *gal* transcription from the P1 promoter only in this population. In vitro kinetic measurements suggest that the stimulation of P1 is due to the product $K_B \times k_f$ being five times greater in the presence of cAMP · CRP than in the absence of the complex (12). It remains to be known how cAMP · CRP affects K_B (closed complex formation) and k_f (isomerization) individually.

The above models of initiation of *gal* transcription are consistent with the 5′-triphosphate analysis of *gal* mRNA in the promoter mutants in vivo as discussed next.

galP211. The *galP211* mutation alters the most conserved second base in the −10 of P1 (Fig. 1, Table 3) and is expected to inactivate this promoter without affecting the activity and cAMP inhibition of P2. The calculated promoter homology scores with *galP211* are essentially unchanged for P2 (59.2) but reduced for P1 (29.6). This is consistent with our observation of the absence of S1-initiated transcript and normal modulation of S2 by cAMP · CRP as was previously shown in vitro (17). These results indicate that the mutation affects open complex formation at P1 both for cAMP-independent and for cAMP-stimulated S1 RNA synthesis.

galP35. The explanation of the behavior of the *galP35* mutation is straightforward: it is a mutation of the cAMP · CRP-binding site (5, 6). Thus, its property should be similar to that of the wild-type promoter in the absence of cAMP. In agreement with this, S2 RNA synthesis from *galP35* in both the presence and absence of cAMP is the same as that from P2 of the wild-type promoter in the absence of cAMP. In both *cya*$^+$ and *cya* strains, little S1 RNA is made for the *galP35* mutant. Thus, failure to bind cAMP · CRP abolished cAMP-dependent stimulation of P1. However, the amount of P1 transcription from this mutant is also reduced threefold compared to that of wild-type P1 in the absence of cAMP. This is very likely because the *galP35* mutation reduces the homology at −35 of P1 (GTCAGA versus GTCACA for *galP*$^+$), thus affecting the cAMP-independent formation of a closed complex at P1 (Table 3). The corresponding *galP35* promoter homology scores are 35.5 for P1 and 58.6 for P2. Ponnambalam et al. (19) also have reported a twofold reduction in cAMP · CRP-independent transcription from P1 or P2 when the −35 region is deleted.

galP18. The single-base-pair deletion of *galP18* eliminates S2 RNA synthesis and causes S1 RNA initiation to be regulated by cAMP · CRP as if it were S2. These in vivo results correlate with those obtained in vitro by Busby et al. (5), who suggested that the *galP18* deletion at −18 caused a fusion of the −35 region of P2 to the −10 region of P1. In the absence of cAMP · CRP, RNA polymerase at the proposed hybrid promoter would preclude the RNA starts at S2. When cAMP · CRP bound at −40, it would prevent RNA polymerase access to the −35 of P2 and thus inhibit S1 starts from the proposed hybrid promoter.

Although the observed behavior of *galP18* can be explained by this proposed hybrid promoter, comparison with the consensus *E. coli* promoter sequences suggests other factors may be involved. First, the spacing between the −35 (of P2) and −10 (of P1) regions for the proposed hybrid promoter in *galP18* is 22 bp (Table 3). This is greater than for any known promoter in the survey by Hawley and McClure (11) and is not recognized as a promoter in the computer program developed by Mulligan et al. (16). Second, for P2 of *galP18* the −35 (of P2) to −10 (of P2) spacing is reduced from 18 (*galP*$^+$) to an ideal 17 bp. This should improve P2 activity, and the promoter homology score of 60.4 is indeed slightly higher than for *galP*$^+$ (59.8). Nonetheless, S2 RNA is detected neither in the absence nor in the presence of cAMP · CRP. A possible explanation is that the single-base-pair deletion at −18 has destroyed the normal alignments of P2 (identical with the consensus promoter) at −17 (T) and at −15/−14 (T-G), which may be more crucial for P2 activity than the weighted values used in the program of Mulligan et al. (16), thus eliminating S2 RNA synthesis. Third, since the 20-bp spacing for P1 is a better match to the consensus promoter than to the hybrid promoter proposed for *galP18*, we propose that the −35 region as well as the −10 region of the P1 is used for S1 RNA in *galP18*. If P1 is used for S1 starts in both *galP*$^+$ and *galP18* cells, we need to explain how cAMP · CRP stimulates the former and inhibits the latter. We propose that the angular relationship between RNA polymerase contacts with the DNA at −10 and contacts with either the DNA or CRP protein at the −35 region is critical to the dual effects on P1 and P2 observed with cAMP · CRP. In *galP18* the single-base-pair deletion at −18 must cause a rotational perturbation that precludes stimulation and facilitates inhibition by cAMP · CRP as seen for *galP*$^+$.

In summary, the following interactions will explain the results obtained so far in studying transcription initiation of the *gal* operon. (i) In the absence of cAMP · CRP, RNA polymerase alone forms closed complexes poorly at P1, but better at P2, whereas isomerization at P1 is more efficient than that at P2. These disparate properties result, however, in about equal promoter strengths for P1 and P2. (ii) cAMP · CRP competes with RNA polymerase to bind to the −40 region, which is critical for binding of the latter to P2. Thus, cAMP · CRP inhibits P2 activity by precluding closed complex formation. cAMP · CRP enhances the activity of P1, possibly by enhancing the isomerization of RNA polymerase at this promoter. (iii) The degree of cAMP · CRP activation of P1 and inhibition of P2 depends on the effective concentration of the cAMP · CRP in the cell. In wild-type cells, growing under conditions of little catabolite repression, both the enhancement of P1 and the inhibition of P2 are about twofold. This modulation of P1 and P2 in opposite directions could be made more severe by exogenous supply of cAMP or by mutation. Another protein:DNA interaction that may prove to be important for *gal* regulation has been reported by Shanblatt and Revzin (20, 21). When RNA polymerase plus cAMP · CRP occupy P1, a second CRP complex can form immediately upstream in the −60 region of the *gal* promoter. The physiological significance of this complex is not known.

ACKNOWLEDGMENTS

We thank S. Venkatesan for the gift of the capping enzyme, S. Busby for the *galP35* and *galP18* mutation-carrying plasmids, D. Court for λ *gal*8 DNA filters, J. Beckwith for strain CA8306, S. Garges for critical reading, and A. Gaddis for typing the manuscript.

This work has been supported in part by the National Cancer Institute (under contract no. N01-C0-75380 with Litton Bionetics, Inc.) and by Public Health Service Biomedical Research Support Award S07RR07160 from the National Institutes of Health to the University of South Carolina.

LITERATURE CITED

1. **Adhya, S., and H. Echols.** 1966. Glucose effect and the galactose enzymes of *Escherichia coli*: correlation between glucose inhibition of induction and inducer transport. J. Bacteriol. **92:** 601–608.
2. **Adhya, S., and W. Miller.** 1979. Modulation of the two promoters of the galactose operon of *Escherichia coli*. Nature (London) **279:**492–494.
3. **Adhya, S. L., and J. A. Shapiro.** 1969. The galactose operon of *E. coli* K-12. Structural and pleiotropic mutations of the operon. Genetics **62:**231–247.
4. **Aiba, H., S. Adhya, and B. de Crombrugghe.** 1981. Evidence for two functional *gal* promoters in intact *Escherichia coli* cells. J. Biol. Chem. **256:**11905–11910.
5. **Busby, S., H. Aiba, and B. de Crombrugghe.** 1982. Mutations in the *Escherichia coli* operon that define two promoters and the binding site of the cyclic AMP receptor protein. J. Mol. Biol. **154:**211–227.
6. **Busby, S., M. Irani, and B. de Crombrugghe.** 1982. Isolation of mutant promoters in the *Escherichia coli* galactose operon using local mutagenesis on cloned DNA fragments. J. Mol. Biol. **154:**197–209.
7. **Chen, C. W., and C. A. Thomas, Jr.** 1980. Recovery of DNA segments from agarose gels. Anal. Biochem. **101:**339–341.
8. **DiLauro, R., T. Taniguchi, R. Musso, and B. de Crombrugghe.** 1979. Unusual location and function of the operator in the *Escherichia coli* galactose operon. Nature (London) **279:**494–500.
9. **Ebright, R., P. Cossart, B. Gicquel-Sanezy, and J. Beckwith.** 1984. Mutations that alter the DNA sequence specificity of the catabolite gene activator protein of *E. coli*. Nature (London)

311:232–235.
10. **Feiss, M., S. Adhya, and D. Court.** 1972. Isolation of plaque forming, galactose-transducing strains of phage lambda. Genetics **171**:180–206.
11. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
12. **Herbert, M., A. Kolb, and H. Buc.** 1986. Overlapping promoters and their control in *Escherichia coli*: the *gal* case. Proc. Natl. Acad. Sci. USA **83**:2807–2811.
13. **Hopkins, J. D.** 1974. A new class of promoter mutations in the lactose operon of *Escherichia coli*. J. Mol. Biol. **87**:715–724.
14. **Irani, M., L. Orosz, S. Busby, T. Taniguchi, and S. Adhya.** 1983. Cyclic AMP dependent constitutive expression of *gal* operon. Use of repressor titration to isolate operator mutations. Proc. Natl. Acad. Sci. USA **80**:4775–4779.
15. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**: 499–560.
16. **Mulligan, M. E., D. K. Hawley, R. Entriken, and W. R. McClure.** 1984. *Escherichia coli* promoter sequences predict *in vitro* RNA polymerase selectivity. Nucleic Acids Res. **12**: 789–800.
17. **Musso, R. E., R. DiLauro, S. Adhya, and B. de Crombrugghe.** 1977. Dual control for transcription of the galactose operon by cyclic AMP and its receptor protein at two interspersed promoters. Cell **12**:847–854.
18. **Nissley, S. P., W. B. Anderson, M. E. Gottesman, R. L. Perlman, and I. Pastan.** 1971. *In vitro* transcription of the gal operon requires cyclic adenosine monophosphate and cyclic adenosine monophosphate receptor protein. J. Biol. Chem. **246**:4671–4678.
19. **Ponnambalam, S., C. Webster, A. Bingham, and S. Busby.** 1986. Transcription initiation at the *Escherichia coli* galactose operon promoters in the absence of the normal −35 region sequences. J. Biol. Chem. **261**:16043–16048.
20. **Shanblatt, S. H., and A. Revzin.** 1983. Two catabolic activator protein molecules bind to the galactose promoter region of *Escherichia coli* in the presence of RNA polymerase. Proc. Natl. Acad. Sci. USA **80**:1594–1598.
21. **Shanblatt, S. H., and A. Revzin.** 1986. The binding of catabolite activator protein and RNA polymerase to the *Escherichia coli* galactose and lactose promoters probed by alkylation interference studies. J. Biol. Chem. **261**:10885–10890.
22. **Shapiro, J. A., and S. L. Adhya.** 1969. The galactose operon of *E. coli* K-12. II. A deletion analysis of operon structure and polarity. Genetics **62**:249–264.
23. **Silverstone, A., R. Ardetti, and B. Magasanik.** 1970. Catabolite-insensitive revertants of *lac* promoter mutants. Proc. Natl. Acad. Sci. USA **66**:773–779.
24. **Summers, W. C.** 1970. A simple method for the extraction of RNA from *Escherichia coli* utilizing diethyl pyrocarbonate. Anal. Biochem. **33**:459–463.
25. **Szybalski, W., H. Kubinski, Z. Hradecna, and W. C. Summers.** 1971. Analytical and preparative separation of the complementary DNA strains. Methods Enzymol. **21**:383–413.
26. **Taniguchi, T., and B. de Crombrugghe.** 1983. Interactions of RNA polymerase and the cyclic AMP receptor protein on the DNA of the *E. coli* galactose operon. Nucleic Acids Res. **11**:5165–5170.
27. **Taniguchi, T., M. O'Neill, and B. de Crombrugghe.** 1979. Interaction site of *Escherichia coli* cyclic AMP receptor protein on DNA of galactose operon promoters. Proc. Natl. Acad. Sci. USA **76**:5090–5094.
28. **Venkatesan, S., A. Gershowitz, and B. Moss.** 1980. Modification of the 5′ end of mRNA. J. Biol. Chem. **255**:903–908.
29. **Venkatesan, S., and B. Moss.** 1981. In vitro transcription of the inverted terminal repetition of the vaccinia virus genome: correspondence of initiation and cap sites. J. Virol. **37**:738–747.
30. **Yamamoto, T., M. E. Sobel, S. L. Adams, V. E. Avvedimento, R. diLauro, I. Pastan, B. de Crombrugghe, A. Showalter, D. Pesciotta, P. Fietzek, and B. Olsen.** 1980. Construction of a recombinant bacterial plasmid containing Pro-α1(I) collagen DNA sequences. J. Biol. Chem. **255**:2612–2615.