Proc. Nat. Acad. Sci. USA
Vol. 72, No. 11, pp. 4394–4398, November 1975
Biochemistry

Initiation of *in vitro* mRNA synthesis from the wild-type *lac* promoter

(regulation of catabolite gene activator protein/repressor function/dinucleotides/restriction fragment)

JOHN MAJORS

Committee on Biophysics, The Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138

Communicated by Charles Yanofsky, September 5, 1975

ABSTRACT **An *in vitro* transcription system, dependent on catabolite gene activator protein (CAP), utilizing a 200 base-pair restriction fragment, has been used to show that the initiation site of the wild-type *Escherichia coli lac* mRNA, and that of two mutants, 8d and p^s, are identical to that previously reported for the CAP-independent promoter mutant UV5. Order of addition experiments are used to show that the binding of *lac* repressor to the operator is competitive with that of the RNA polymerase (nucleosidetriphosphate:RNA nucleotidyltransferase, EC 2.7.7.6) to the promoter, thus demonstrating functional overlap of the operator and promoter sites.**

Control of gene expression in prokaryotic cells is exercised primarily at the level of transcription. In the lactose operon of *Escherichia coli*, a model system for the study of this control, coordinate transcription by the RNA polymerase of the three genes responsible for lactose metabolism is subject to the negative control of the lactose repressor protein and the positive control of the catabolite gene activator protein, CAP. In the absence of a suitable inducer, the *lac* repressor binds to the operator and prevents expression of the *lac z*, *y*, and *a* genes (1, 2). CAP is the cAMP-activated mediator of catabolite repression in *E. coli* (3, 4), a mechanism that enables a cell growing in the presence of glucose to restrict synthesis of enzymes responsible for the less efficient metabolism of certain other sugars (5). A low extracellular glucose concentration induces a high intracellular level of cAMP (6), whereby CAP is activated and potentiates initiation by the RNA polymerase at promoters of catabolite repressible genes (7–9).

In vitro transcription from the *lac* promoter has been shown to be sensitive to these controls and to promoter mutations in the template DNA (9, 10). The sequence of the first 63 bases of the mRNA from the CAP-independent promoter mutant, UV5, has been determined (11), as has that of the entire promoter-operator region of the wild-type strain (12). The operator has been isolated; its base sequence has been determined and contains the initiation site for the UV5 mRNA (13). This evidence argues for a simple competitive binding model for the interaction between *lac* repressor, the RNA polymerase, and the UV5 control region. Earlier results by Chen *et al.* (14) suggested a different model for wild-type *lac* control in which the repressor and polymerase binding are completely independent, presenting the possibility that the wild-type mRNA initiation site differs from that of the mutant UV5. I report here experiments which show that the wild-type mRNA initiation site is in fact identical to that of UV5, and demonstrate that, contrary to the results of Chen *et al.*, the binding of repressor to the operator and the binding of polymerase to the promoter are mutually exclusive events.

Abbreviations: CAP, catabolite gene activator protein; TBM, 0.1 M Tris-borate, pH 8.3, 2 mM $MgCl_2$; TBE, 0.1 M Tris-borate, pH 8.3, 1 mM EDTA; IPTG, isopropyl-β-D-thiogalactopyrannoside.

METHODS

The *Hae* 200 DNA Fragment. The two-hundred base pair *Hae* 200 fragment, which carries the *lac* promoter-operator region, was isolated starting from intact λ*plac*$_5$ phage. Sonicated, thousand base pair fragments were prepared and purified for *lac*-containing DNA as described by Gilbert and Maxam (13). These fragments, typically 500 μg in a volume of 2 ml, were digested for 1 hr at 37° with *Hae*III enzyme in a buffer containing 10 mM Tris·HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM dithiothreitol, and 200 μg/ml of bovine serum albumin. The *Hae*III enzyme, purified by the method of Middleton *et al.* (15), was kindly supplied by Edith Butler. The cut fragments were then made up to 2 M NH_4OAc, precipitated with 2 volumes 95% ethanol, resuspended in 0.5 M NH_4OAc, again precipitated with ethanol, and then electrophoresed through a 5% polyacrylamide TBM (0.1 M Tris-borate, pH 8.3, 2 mM $MgCl_2$) gel (13). The *Hae* 200 fragment, visualized with ultraviolet light by viewing the gel against a fluorescent screen, was cut out, eluted overnight into 0.5 M NH_4OAc, 0.1% sodium dodecyl sulfate, 10 mM $Mg(OAc)_2$, and 0.1 mM EDTA, then precipitated twice with ethanol and used in the *in vitro* reaction described below.

***In Vitro* Synthesis of Wild-Type mRNA.** Typical reaction mixtures contained in a volume of 25 μl: 30 mM Tris·HCl, pH 8.0, 3 mM $MgCl_2$, 100 mM KCl, 0.1 mM EDTA, 0.1 mM dithiothreitol, 0.5 mg/ml of bovine serum albumin, 1 mM cAMP, 0.1 μg of *Hae* 200, 1 μg of RNA polymerase (nucleosidetriphosphate:RNA nucleotidyltransferase, EC 2.7.7.6), and 0.25 μg of CAP. RNA polymerase was purified by the method of Berg *et al.* (16) through the glycerol gradient step and was stored at 4° as it came off the gradient. CAP was purified by the method of Pastan *et al.* (17) with the modification that dithiothreitol was excluded throughout. In a typical reaction, salts, DNA, cAMP, and CAP were preincubated at 37°; after 3 min RNA polymerase was added and incubation was continued for an additional 5–10 min. At this time heparin(Upjohn), a polyanionic inhibitor of free RNA polymerase (18), was added to a concentration of 100 μg/ml to eliminate nonspecifically bound polymerases. One minute later synthesis was begun with the addition of triphosphates, ATP and GTP to a concentration of 200 μM, and UTP and CTP to a concentration of 2–10 μM. Either or both of the latter were [α-^{32}P]NTP at a specific activity of greater than 100 Ci/mmol (New England Nuclear). After 10 min, synthesis was terminated by the addition of 0.5 ml of an ice-cold solution containing 2 M NH_4OAc, 2 mM EDTA, and 100 μg/ml of tRNA (Schwartz). Two volumes of 95% ethanol were added, and the RNA was precipi-

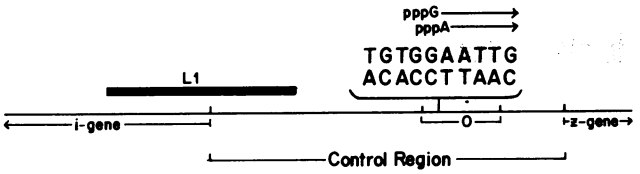

FIG. 1. The *lac Hae* 200 fragment with significant genetically and biochemically defined regions as determined by Gilbert *et al.* (20, 33) and Beckwith *et al.* (21, 22). The sequence to the left of the initiation site was determined by both Dickson *et al.* (12) and A. Maxam (unpublished results).

tated. The precipitate was resuspended in 0.5 ml of 0.5 M NH_4OAc, again precipitated with ethanol, resuspended in 100 μl of distilled water, and lyophilized. The RNA was then electrophoresed through a 12% polyacrylamide, 7 M urea, TBE (0.1 M Tris-borate, pH 8.3, 1 mM EDTA) gel (13). The RNA band, visualized by autoradiography, was cut out and its radioactivity was determined in a scintillation vial or it was eluted into 0.5 ml of the elution buffer described in the previous section containing 20 μg of tRNA.

Alkaline Phosphatase Digestion. RNA was synthesized in reaction volumes of 50 μl. After 10 min of synthesis at 37° the mixture was divided into equal aliquots, to one of which was added 20 μg of bacterial alkaline phosphatase (Worthington), and incubation was continued at 37° for 1 hr. The reaction was terminated with the addition of 0.5 ml of 2 M NH_4OAc, 2 mM EDTA, 100 μg/ml of tRNA, and the sample was prepared for electrophoresis as described previously.

Fingerprinting. Fingerprinting techniques are described in Barrel (19) and in Maxam and Gilbert (13). The first dimension buffer was 5% HOAc, 0.1% pyridine, 7 M urea, and 5 mM EDTA. The high EDTA concentration was to prevent any triphosphate end from smearing. The second dimension buffer was 7% formic acid, pH 1.7.

RESULTS

These experiments were made possible by the recent isolation of a small promoter-containing DNA fragment, generated by the digestion of thousand base pair sonicated *lac* fragments with the *Hae*III restriction endonuclease of *Hemophilus aegyptius* (20). The fragment, detailed in Fig. 1 and designated *Hae* 200, is about 200 base pairs long and was originally shown to contain the complete *lac* control region by the following criteria: (*i*) In an *in vitro* system using the *Hae* 200 fragment from the UV5 mutant and the dinucleotide GpA as a primer, UV5 mRNA was synthesized whose sequence extends into the region of the message coding for the NH_2-terminus of β-galactosidase, the product of the *z* gene (N. Maizels, unpublished results). (*ii*) In a digest of DNA from a strain containing the L1 deletion, known to cover both the COOH-terminus of the *i* gene and part of the *lac* promoter (21, 22), only the *Hae* 200 fragment showed an altered mobility on polyacrylamide gels (W. Gilbert, J. Gralla, and A. Maxam, to be published); the deletion falls completely within this fragment. The *Hae* 200 piece therefore must contain the entire region from the end of the *i* gene to the beginning of the *z* gene.

The fragment was isolated in microgram quantities starting from 50 or 100 mg of sonicated $\lambda plac_5$ fragments and used as a template in the *in vitro* transcription system described in *Methods*. (Although the majority of *lac* specific RNA synthesized *in vitro* from whole $\lambda plac_5$ DNA originates from a phage promoter and is not subject to either repressor or CAP control, the strain does contain a wild-type *lac* promoter and operator which can be isolated apart from the phage promoter on the *Hae* 200 fragment.) Fig. 2 shows that the major synthesis product is a broad band on a 12% polyacrylamide, 7 M urea TBE gel, with a mobility corre-

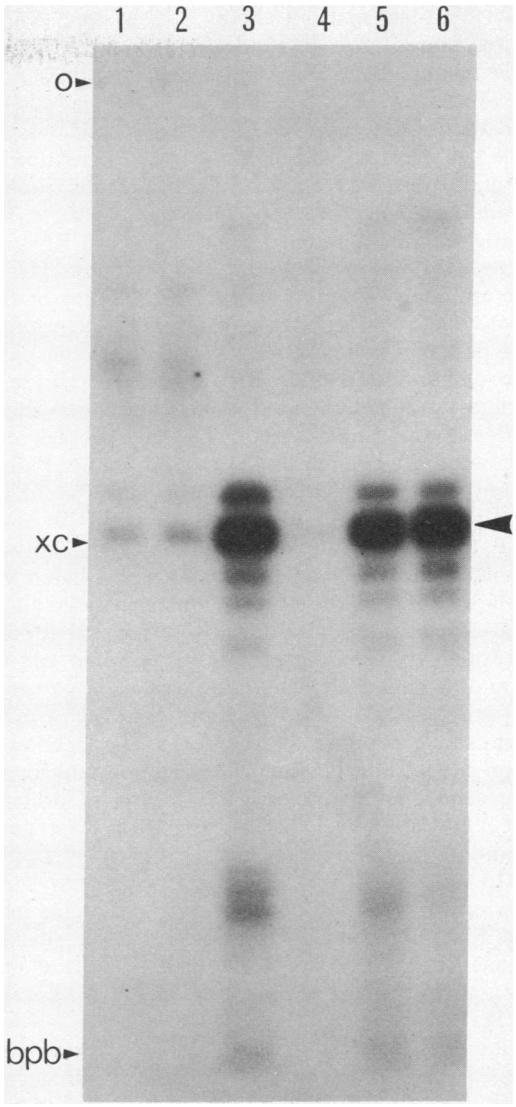

FIG. 2. Polyacrylamide gel electrophoresis pattern of RNA synthesized *in vitro* from the wild-type *lac Hae* 200 fragment. In each case the major band (◀), about 56 bases long, was cut out, placed in a scintillation vial, and counted using Cerenkov radiation. Reaction conditions for each were as described in *Methods* with the indicated modifications:

	Cerenkov cpm
(*i*) − cAMP + CAP	814
(*ii*) + cAMP − CAP	1,072
(*iii*) + cAMP + CAP	18,262
(*iv*) + cAMP + CAP + 0.25 mg of *lac* repressor added before RNA polymerase	384
(*v*) + cAMP + CAP + 0.25 mg of *lac* repressor + 1 mM IPTG, an inducer of the repressor	11,300
(*vi*) + cAMP + CAP + 0.25 mg of *lac* repressor added 10 min after RNA polymerase	15,262

o is the origin; xc is the position of the xylene cyanol dye marker; and bpb is the position of the bromphenol blue dye marker. IPTG, isopropyl-β-D-thiogalactopyrannoside.

sponding to a piece of RNA about 56 bases long. As can be seen in the same figure, its synthesis (and that of a number of minor bands, about which more will be said shortly) is about 20-fold stimulated by cAMP and CAP and is sensitive to the action of *lac* repressor. The major band was eluted from the gel, digested with either T1 or pancreatic ribonuclease, and subjected to two-dimensional electrophoresis. The resultant fingerprints contained each of those T1 and pancreatic fragments found by N. Maizels in RNA of the same length synthesized from the UV5 promoter (11), with a possible ambiguity existing only in the identity of the triphosphate end. This identity was indirectly determined by treating part of a reaction mixture with alkaline phosphatase. Treated and untreated aliquots were then run in parallel through a 12% gel, digested with either pancreatic or T1 ribonuclease, and fingerprinted. The appearance of additional label in a fragment of the treated material identifies that fragment as the 5′-triphosphate end. In the T1 fingerprint of RNA labeled with both [α-^{32}P]UTP and [α-^{32}P]CTP, the treated sample has additional label only in ApApUpUpGp and is missing a spot that moves like a triphosphate end in a fingerprint of the nonphosphatased aliquot. A pancreatic fingerprint of identical RNA shows that phosphatase treatment transfers label only to ApApUp and GpApApUp and nowhere else. Therefore, the message can begin either pppApApUpUpGp . . . or pppGpApApUpUpGp . . . , the same result as found for UV5 (11). The location of the initiation site within the *Hae* 200 fragment and in relation to other genetic landmarks is shown in Fig. 1. Although the ratio of ATP starts to GTP starts varies with the experimental conditions, approximately equal recovery of pppApApUpUpGp and ApApUpUpGp in the T1 fingerprints of RNA synthesized in a number of different experiments indicates that the RNA polymerase will initiate at either site with almost equal probability, in contrast to the results of synthesis off UV5 at 25°, where a predominance of ATP starts was seen (11).

A number of the minor bands, whose syntheses are subject to both CAP and repressor control, were also fingerprinted. Each of those examined initiated at the same site as the major product. The longest was 63 bases long (see Fig. 2) and contained RNA complementary to the entire right end of the *Hae* 200 fragment. Therefore, under the conditions of the synthesis a majority of the polymerases halt before reaching the end of the fragment. This phenomenon is enhanced by the use of heparin; more polymerases reach the end when rifampicin is used or when heparin is omitted. This pausing site does not correspond to any of those reported by Maizels, and is a function of the polymerase having to "run-off" the end of the DNA fragment.

This same system has been used to synthesize mRNA from the *Hae* 200 fragments of several promoter mutants: UV5, as discussed earlier; 8d, a UV5-like partially CAP-independent mutant (23); and p^s, a mutant with a greater sensitivity to CAP in *in vitro* systems (24). The RNA polymerase initiates from all of these promoters at the same site, again showing no preference for ATP or GTP as the initiating triphosphate.

Dinucleotides may be used to prime synthesis by RNA polymerase when the purine triphosphate concentrations are too low to allow initiations to occur (25, 26). The ability of the polymerase to initiate synthesis at sites close to the above reported initiation site was checked by using primer dinucleotides complementary to the known sequence around this site (see Fig. 1). The dinucleotides UpG, GpU, GpG, GpA, ApA, and ApU were each tested at concentrations of 200 μM with the triphosphates all held below 10 μM. Of these, only GpA, GpG, and ApA stimulated any mRNA synthesis; GpA was about 3-fold more efficient than GpG and many-fold more efficient than ApA, which barely promotes synthesis.

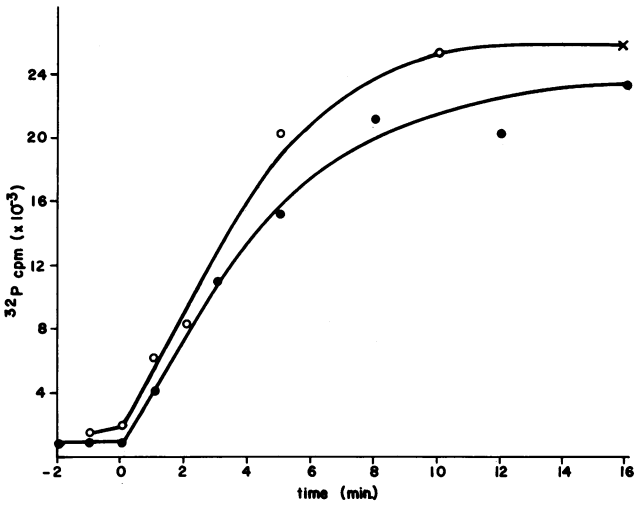

FIG. 3. Functional overlap of the repressor and polymerase interaction sites demonstrated by measuring the kinetics of formation of a heparin-resistant complex. In the first experiment (●) RNA polymerase was added to a reaction mixture containing CAP, cAMP, and *Hae* 200 at 0 min. *lac* repressor was added at the indicated times. Five minutes later heparin was added to a concentration of 100 μg/ml, followed 1 min later by the addition of triphosphates, including [α-^{32}P]UTP. The synthesized RNA was electrophoresed on a 12% polyacrylamide gel and the major band cut out and counted as described in *Methods*. In the second experiment (○) repressor was added to a reaction mixture containing CAP, cAMP, and *Hae* 200. After 5 min at 37° RNA polymerase was added. Ten minutes later, at 0 min on the time scale in the figure, the repressor was induced by the addition of IPTG to a concentration of 1 mM. Heparin to 100 μg/ml was added at the indicated points in the time course, and synthesis was begun 1 min later with the addition of triphosphates. (×) is the level of RNA synthesis when IPTG is added before the addition of *lac* repressor.

How the *lac* Repressor Works. The polymerase initiates synthesis of the wild-type mRNA within the operator, functionally defined as that sequence of DNA made resistant to DNase digestion by the interaction of the *lac* repressor. It follows that the repressor probably functions by preventing the polymerase from forming an initiation complex. This scheme is contrary to the accepted model for the repressor-polymerase interaction with the wild-type *lac* promoter (14), though consistent with that for both UV5 and p^s (10, 14). To check this prediction, I did simple order of addition experiments: adding repressor and polymerase at different times to a reaction mixture containing DNA and CAP. The major synthesis product was isolated on polyacrylamide gels, cut out, and counted for incorporated radioactivity. Fig. 3 shows that within an incubation time of 8 min a large number of RNA polymerase–*lac* promoter complexes become refractory to the influence of *lac* repressor. Presumably this lag represents the time it takes the polymerase to attach itself to the promoter and form a heparin-resistant complex, since a similar lag is seen if one alters the above experiment by adding heparin instead of repressor at the indicated times and then allows synthesis to proceed (data not shown).

In a different experiment repressor, CAP, and *Hae* 200 were preincubated at 37° before the addition of RNA polymerase. Polymerase was then added, and incubation contin-

ued for another 10 min, at which time repressor was induced by the addition of IPTG. The ability of the polymerase to form a heparin-resistant initiation complex after induction was then monitored by adding heparin at different times with respect to this induction. As can also be seen in Fig. 3, the kinetics of this formation were essentially identical to those found in the previous experiment, requiring upwards of 8 min after induction for full expression to be reached. Because IPTG-induced repressor dissociates from the operator extremely rapidly (27), this lag again probably represents the association of free RNA polymerases. A similar lag has been reported for the association of the polymerase with the *trp* promoter (28) and with the *lac* p^s promoter (9). Hinkle and Chamberlin (28), using filter binding, have shown that RNA polymerase forms a tight complex with T7 DNA with a half-life of on the order of 20 sec, while the above-mentioned systems show half-life association times of on the order of 2–4 min. The reason for the difference between these association times is not known, although in some cases a small amount of core enzyme in the polymerase preparation might partially inhibit the formation of initiation complexes by the holoenzyme.

Because the results of Chen *et al.* were obtained from experiments in which rifampicin rather than heparin was used as an inhibitor of free RNA polymerases, I have also done order of addition experiments using this inhibitor, adding it to a concentration of 10 μg/ml along with the triphosphates. There was no significant difference between these results and those with heparin except for the previously mentioned difference in the pausing pattern.

DISCUSSION

The use of restriction endonucleases to generate small, promoter-containing DNA fragments greatly facilitates the study of the interactions of RNA polymerase and related regulatory proteins with these promoters. RNA synthesized in an *in vitro* system using such a fragment as a template can be limited solely to that initiated from the promoter of interest. The ease with which the sequences of the fragment and the RNA synthesized from it may be determined then allows one to make correlations between sequence information and other experimentally determined parameters. The *lac Hae* 200 fragment has been used in these experiments to show that the wild-type mRNA and that from a number of different mutant promoters are all initiated at the same site. Each of these mutations then must represent either an alteration in the affinity of the polymerase for an already existing binding site or an altered ability of the polymerase to approach this site, rather than the substitution of a completely new promoter, as was postulated by Chen *et al.* (14).

The ambiguity in the initiating triphosphate of the *lac* mRNA reported here and elsewhere (11) implies a flexibility on the part of the polymerase in the initiation process. The sequences of a number of different RNA polymerase binding sites have been reported (see ref. 29); although a region of homology has been found about 10 base pairs to the left of the mRNA initiation sites (29), there is no discernible homology at the initiation site itself. This supports a model in which the polymerase is held in the correct conformation for initiation by interactions removed from the actual initiation site, which then becomes any properly situated purine-coding base pair.

The simplest and most efficient way for a repressor to work would be for it to block the approach of the polymerase to the binding site and prevent the formation of an initiation complex. Evidence for this type of control has been reported for the *gal* operon (30), the *trp* operon (28), the early lambda promoter, p_r (31), and for CAP-independent mutants of *lac*. It is not surprising that the same is true for wild-type *lac*, although an earlier report (14), using hybridization assays to monitor synthesis from whole λ*h80dlac* DNA, argued for a model in which repressor and polymerase binding were completely independent. The reason for the discrepancy between that report and these results remains obscure. The most likely explanation for the functional overlap reported here is that *lac* repressor physically covers those sequences with which the polymerase must interact. It is known that DNA fragments protected from DNase digestion by polymerase bound to both the *lac* and lambda promoters contain substantial operator sequence information, though on *lac* these sequences are on the mRNA proximal end of the fragment (J. Gralla, to be published) while in lambda they are on the mRNA distal end (32). The system described here is being used to investigate in more detail the interactions between CAP, polymerase, repressor, and the wild-type *lac* promoter.

1. Jacob, F. & Monod, J. (1961) *J. Mol. Biol.* **3**, 318–356.
2. Gilbert, W. & Muller-Hill, B. (1967) *Proc. Nat. Acad. Sci. USA* **58**, 2415–2421.
3. Zubay, G., Schwartz, D. & Beckwith, J. (1970) *Proc. Nat. Acad. Sci. USA* **66**, 104–110.
4. Emmer, M., deCrombrugghe, B., Pastan, I. & Perlman, R. (1970) *Proc. Nat. Acad. Sci. USA* **66**, 480–487.
5. Magasanik, B. (1970) in *The Lac Operon*, eds. Zipser, D. & Beckwith, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), pp. 189–219.
6. Makman, R. & Sutherland, E. (1965) *J. Biol. Chem.* **240**, 1309–1314.
7. Silverstone, A., Arditti, R. & Magasanik, B. (1970) *Proc. Nat. Acad. Sci. USA* **66**, 773–779.
8. deCrumbrugghe, B., Chen, B., Anderson, W., Gottesman, M., Perlman, R. & Pastan, I. (1971) *J. Biol. Chem.* **246**, 7343–7348.
9. Nissley, S., Anderson, W., Gottesman, M., Perlman, R. & Pastan, I. (1971) *J. Biol. Chem.* **246**, 4671–4678.
10. Eron, L. & Block, R. (1971) *Proc. Nat. Acad. Sci: USA* **68**, 1828–1832.
11. Maizels, N. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 3585–3589.
12. Dickson, R., Abelson, J., Barnes, W. & Reznikoff, W. (1975) *Science* **187**, 27–35.
13. Gilbert, W. & Maxam, A. (1973) *Proc. Nat. Acad Sci. USA* **70**, 3581–3584.
14. Chen, B., deCrumbrugghe, B., Anderson, W., Gottesman, M., Pastan, I. & Perlman, R. (1971) *Nature New Biol.* **233**, 67–70.
15. Middleton, J., Edgell, M. & Hutchison C., III, (1972) *J. Virol.* **10**, 42–50.
16. Berg, P., Barret, K. & Chamberlin, M. (1971) in *Methods in Enzymology*, eds. Grossman, L. & Moldave, K., (Academic Press, New York), Vol. 21D, pp. 506–509.
17. Pastan, I., Gallo, M. & Anderson, W. (1974) in *Methods in Enzymology*, eds. Hardman, J. & O'Malley, B. (Academic Press, New York), Vol. 28, pp. 367–376.
18. Zillig, W., Zechel, K., Radussay, D., Schachner, M., Sethi, V., Palm, P., Heil, A. & Seifert, W. (1970) *Cold Spring Harbor Symp. Quant. Biol.* **35**, 47–58.
19. Barrel, B. (1971) in *Procedures in Nucleic Acid Research*, eds. Cantoni, G. & Davies, D. (Harper & Row, New York), Vol. 2, pp. 751–779.
20. Gilbert, W., Gralla, J., Majors, J. & Maxam, A. (1975) in *Protein-Ligand Interactions*, eds. Sund, H. & Blauer, G. (Walter deGruyter & Co., Berlin), pp. 193–210.

21. Scaife, J. & Beckwith, J. (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 403–408.
22. Miller, J., Ippen, K., Scaife, J. & Beckwith, J. (1968) *J. Mol. Biol.* **38**, 413–420.
23. Arditti, R., Grodzicker, T. & Beckwith, J. (1973) *J. Bacteriol.* **114**, 652–655.
24. Downey, K., Jurmar, B. & So, A. (1971) *Biochemistry* **10**, 4970–4975.
25. Minkley, E. & Pribnow, D. (1973) *J. Mol. Biol.* **77**, 255–277.
26. Riggs, A., Newby, R. & Bourgeois, S. (1970) *J. Mol. Biol.* **51**, 303–314.
27. Squires, C., Lee, F. & Yanofsky, C. (1975) *J. Mol. Biol.* **91**, 93–111.
28. Hinkle, D. & Chamberlin, M. (1972) *J. Mol. Biol.* **70**, 187–195.
29. Pribnow, D. (1975) *Proc. Nat. Acad. Sci. USA* **72**, 784–788.
30. Nakanishi, S., Adhya, S., Gottesman, M. & Pastan, I. (1973) *J. Biol. Chem.* **248**, 5937–5942.
31. Pirotta, V. (1975) *Nature* **254**, 114–117.
32. Walz, A. & Pirotta, V. (1975) *Nature* **254**, 118–121.
33. Gilbert, W., Maizels, M. & Maxam, A. (1973) *Cold Spring Harbor Symp. Quant. Biol.* **38**, 845–855.