Transcription of Bacteriophage λ DNA in Vitro Using Purine Nucleoside 5′-[γ-S]Triphosphates as Affinity Probes for RNA Chain Initiation†

M. Mitchell Smith,*,‡ Anthony E. Reeve, and Ru Chih C. Huang

ABSTRACT: Bacteriophage λcb2 DNA was transcribed with *Escherichia coli* RNA polymerase in vitro using either adenine or guanosine 5′-[γ-S]triphosphate as one of the nucleotide substrates. Transcripts bearing either [S]pppAp or [S]pppGp at their 5′ end were isolated by affinity chromatography on mercury-agarose and the known in vitro transcription map of bacteriophage λ was used to test the accuracy of initiation obtained when using these analogues. Two RNA products originating from the λ promoters p_L and p_O (p_L') were examined. Of the four known transcripts synthesized on λcb2 DNA in vitro, three initiate with ATP and one initiates with GTP. Since the major leftward RNA originating at p_L is the only leftward transcript to begin with ATP, this first sequence was isolated by binding RNA which initiated with [S]pppAp to mercury-agarose followed by hybridization of the bound RNA to the separated l-strand of λ DNA. This RNA gave an oligonucleotide fingerprint which identified it as the major leftward RNA and showed that the transcript initiated at or near its correct startpoint with the sequence [S]pppApU. The only RNA known to initiate with GTP in vitro is the minor leftward 4S RNA, or *oop* RNA, transcribed from p_O. Therefore, this product was purified directly by simply binding RNA which initiated with [S]pppGp to mercury-agarose. This second transcript was proven by fingerprint analysis to be *oop* RNA and it initiated at its correct startpoint with the sequence [S]pppGpUpU. The relative efficiency of transcription from the λ promoters was unaffected by replacement of ATP or GTP with the sulfur-containing nucleotides, and no errors in the fidelity of transcription were detected as a result of the substitution.

The control of initiation of specific RNA sequences is an important element in gene regulation. The ability to initiate RNA transcripts correctly provides a critical test of any in vitro system designed to identify regulatory determinants of gene expression. Yet, the detailed analysis of RNA initiation, particularly of rare sequences, is difficult. Even in the simpler bacteriophages, identification of a primary transcript and its initiating nucleotide is an intricate problem (Chamberlin, 1974a,b). In the eukaryotes the difficulties are much greater. With a few exceptions, such as 5S ribosomal RNA (Hatlen et al., 1969), not even the size of the primary transcript of a given eukaryotic RNA is known, much less its initiating nucleotide sequence. Consequently, most studies to date have been restricted to examination of the initiation of the abundant low molecular weight RNA species transcribed by polymerase III, particularly 5S rRNA (Marzluff et al., 1974).

Detection of initiation in vitro has for the most part utilized nucleoside 5′-[γ-^{32}P]triphosphate incorporation into RNA with the retention of a radiolabeled γ-phosphate at the 5′ end (Maitra & Hurwitz, 1965). This method results in the incorporation of a single radioactive molecule per initiated transcript and serves well for abundant RNA classes. For rare messenger RNA transcripts from single copy genes, however, even high specific activity γ-^{32}P-labeled nucleotides fail to give sufficient sensitivity to make the assay practical.

With these problems in mind, we have undertaken the development of an affinity binding assay for RNA initiation. The method utilizes analogues of the nucleoside triphosphates containing a sulfur on the γ-phosphate (Goody & Eckstein, 1971). Since phosphodiester bond formation requires elimination of the β- and γ-phosphates from the nucleotide precursors, incorporation of purine nucleoside 5′-[γ-S]triphosphates into RNA should result in the retention of sulfur only at the 5′ end of transcripts which initiated with that specific purine. These sulfur-containing transcripts can then be selectively bound to an affinity column of mercury-agarose. This method has the advantage of amplifying the initiation event since radiolabel may be incorporated throughout the length of the molecule instead of only at the 5′-end γ-phosphate. In addition, it allows the physical isolation of in vitro initiated transcripts based on their initiating nucleotide.

In a previous report (Reeve et al., 1977) we demonstrated that purine nucleoside 5′-[γ-S]triphosphates could be incorporated into RNA transcribed by *Escherichia coli* RNA polymerase from the synthetic DNA templates $d(A\text{-}T)_n \cdot d(A\text{-}T)_n$ and $d(G\text{-}C)_n \cdot d(G\text{-}C)_n$. The sulfur was retained exclusively at the 5′ end of initiated transcripts and enabled these molecules to be selectively retained on mercury-agarose. In this paper we wish to report experiments designed to test the ability of the sulfur nucleotides to initiate at defined promoter sites and allow isolation of these specific sequences by affinity chromatography.

For these studies we have used the transcription of λcb2 DNA by *E. coli* RNA polymerase holoenzyme. The promoters, polarities, and initiating sequences of the four RNA products transcribed in vitro from λcb2 DNA are illustrated

† From the Department of Biology, The Johns Hopkins University, Baltimore, Maryland 21218. *Received July 22, 1977.* This research was supported by National Institutes of Health Grants R01CA13953, 1T01HD00826, and 5T01AG00027. M.M.S. was an NIH Postdoctoral Fellow and is currently the recipient of a Helen Hay Whitney Foundation Postdoctoral Fellowship. A.E.R. was the recipient of U.S. Public Health Service International Research Fellowship No. 5F05TW02175-03.

‡ Current address: Department of Molecular Biology, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh, Scotland EH9 3JR.

0006-2960/78/0417-0493$01.00/0

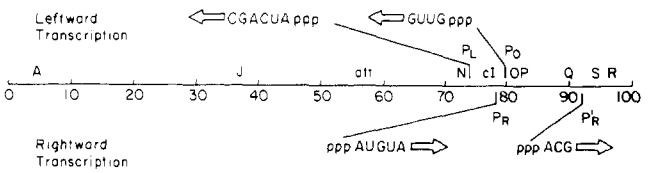

FIGURE 1: In vitro transcription of bacteriophage λ DNA. The points of initiation, polarity, and initial nucleotide sequences of the four predominant RNA products synthesized from λcb2 DNA in vitro are shown. The symbols directly above the physical map, such as A, J, cI, and R, represent genetic markers of phage λ and mark their positions along the scale 0 to 100 dividing the length of the genome. Ths positions of the major leftward and rightward promoters, p_L and p_R, as well as the minor promoters, p_O and p_R', are indicated along with the RNA sequence originating at that promoter. (Adapted from Blattner & Dahlberg, 1972).

in Figure 1 (Blattner & Dahlberg, 1972). The strategy of the experiments is as follows. To study ATP-initiated sequences, RNA is transcribed in the presence of [γ-S]ATP[1] and isolated on mercury–agarose. The bound RNA is then hybridized with separated l-strand DNA, i.e., DNA complementary to leftward transcripts. If the sulfur analogues are initiating accurately, the isolated RNA should be the major leftward RNA originating from the promoter p_L, designated here as p_L RNA. Similarly for GTP-initiated sequences, RNA is transcribed in the presence of [γ-S]GTP. Since the minor leftward RNA, designated *oop* RNA (Hayes & Szybalski, 1973), is the only product initiating with GTP on λcb2 DNA, this sequence should be obtained directly by binding to mercury–agarose.

We report here that these predictions are, in fact, correct. *E. coli* RNA polymerase appears to initiate and transcribe bacteriophage λ DNA accurately using [γ-S]ATP and [γ-S]GTP. These analogues are therefore extremely useful probes for the detection of RNA chain initiation.

Materials and Methods

Reagents and Bacteriophage Strains. Adenosine 5′-[γ-S]triphosphate was purchased from Boehringer Mannheim Biochemicals. Guanosine 5′-[γ-S]triphosphate was generously donated by Dr. W. Waessle of Boehringer. Unlabeled nucleotides were from Sigma Biochemicals, and radiolabeled nucleotides were obtained from both New England Nuclear Corp. and Amersham/Searle Corp. *Escherichia coli* RNA polymerase (EC 2.7.7.6) was obtained from Miles Laboratories. Organomercurial agarose was purchased as Affi-Gel 501 from Bio-Rad Laboratories. Bacteriophage λ strain λcb2 was the gift of Dr. M. Rhoades and strain λb2cI857Sam7 lysate was the generous gift of Dr. E. N. Moudrianakis. Separated strands of λ DNA were prepared from λcI857Sam7 DNA purchased from Bethesda Research Laboratories.

Purification of Nucleoside 5′-[γ-S]Triphosphates. Preparations of [γ-S]ATP and [γ-S]GTP routinely contained 10–20% of the corresponding nucleoside diphosphate. Therefore, the sulfur-containing nucleotides were purified over mercury–agarose prior to their use in transcription reactions. Approximately 4 μmol of sulfur nucleotide in 1 mL of 0.1 M NaCl, 10 mM Tris-HCl (pH 7.9), 1 mM EDTA were applied to Affi-Gel 501 equilibrated in the same buffer. Following binding, the column was washed with deionized water to remove salts. The thiol analogue was then eluted with 10 mM 2-mercaptoethanol in deionized water and lyophilized. It was subsequently adjusted to 10 mM with deionized water and stored frozen at −20 °C.

Preparation of [^{3}H]Nucleoside 5′-[γ-S]Triphosphate Precursors. [8-^{3}H]Adenosine 5′-[γ-S]triphosphate and [8-^{3}H]guanosine 5′-[γ-S]triphosphate were prepared by enzymatic exchange of γ-thiophosphate from [γ-S]ATP to either [^{3}H]ADP or [^{3}H]GDP with nucleoside diphosphate kinase (EC 2.7.4.6) according to Goody et al. (1972). The reaction (0.1 mL) was terminated by diluting to 1 mL with 0.01 M Tris-HCl (pH 7.9), 0.1 M NaCl, 0.01 M EDTA, 0.1% NaDodSO$_4$. The sample was sequentially phenol and ether extracted and then applied to an Affi-Gel 501 mercury–agarose column to bind the sulfur-containing nucleotides. The column was washed with 0.1 M ammonium acetate, 0.01 M EDTA and the tritium radiolabeled thiol nucleotide was displaced with 0.1 M ammonium acetate, 0.01 M dithiothreitol. The product was lyophilized and stored in a 50% ethanol solution at −20 °C.

Transcription of RNA. RNA was transcribed from initiation complexes by pulse synthesis at low temperature as described by Blattner & Dahlberg (1972). For (A-U)$_n$ synthesis, the synthetic template d(A-T)$_n$·d(A-T)$_n$ was used. Bacteriophage λ RNA was transcribed from either λcb2 or λb2cI-857Sam7 DNA. The DNA of λ b2 deletion mutants was used to eliminate initiation from the b2 region (Roberts, 1969).

The preincubation mixture contained in 0.4 mL, 200 μg of DNA, 200 μg of *E. coli* RNA polymerase, 3.2 μmol of EDTA, 1.6 μmol of $MgCl_2$, 20 μmol of Tris-HCl (pH 7.9), and 68 μmol of KCl. Reactions also included 40 pmol of each nucleoside triphosphate. From 1 to 2 mCi of [α-^{32}P]UTP was included as the radiolabeled nucleotide. For synthesis of sulfur-containing RNA, either [γ-S]GTP or [γ-S]ATP completely replaced the corresponding GTP or ATP in the reaction.

This preincubation mixture was incubated at 25 °C for 2 min at which time the polymerization was started by the addition of 0.04 mL of deionized water containing 16 μmol of $MgCl_2$ and 4 μg of rifampin. The reaction was stopped by addition of 0.4 mL of 0.2 M EDTA and chilling on ice. The mixture was brought to 2 mL with 0.1 M sodium acetate buffer (pH 5), 5 mM EDTA, 0.1% NaDodSO$_4$, and extracted 15 min on ice with water-saturated phenol. The aqueous phase was removed and neutralized with 0.2 mL of 1 M Tris-HCl (pH 7.9).

Isolation of [γ-S]RNA on Mercury–Agarose. All steps were performed at room temperature. The neutralized aqueous phase containing the [γ-S]RNA was subjected to gel filtration by chromatography on Bio-Gel P6 (2 × 8 cm) equilibrated with 0.1 M NaCl, 10 mM Tris-HCl (pH 7.9), 1 mM EDTA, 0.1% NaDodSO$_4$ (TNES buffer). The excluded radioactivity was applied directly to an 8-mL column (1.2 × 5 cm) of Affi-Gel 501 mercury-agarose equilibrated in TNES buffer. The column was washed with 40 mL of TNES buffer and the bound [γ-S]RNA was eluted with TNES buffer containing 10 mM dithiothreitol. The eluted [γ-S]RNA was ethanol precipitated along with carrier yeast RNA. The mercury–agarose was regenerated by washing with TNES buffer to remove dithiothreitol followed by 2 volumes of TNES buffer containing 10 mM $HgCl_2$. After extensive washing with TNES buffer, the mercury–agarose was fully regenerated.

While Affi-Gel 501 was used successfully in these studies, primarily because of the short chain length of the pulse synthesized [γ-S]RNA, it must be noted that this mercury–agarose may not be acceptable for other studies. We have recently reported that Affi-Gel 501, and other mercury gels prepared

[1] Abbreviations: [γ-S]ATP, adenosine 5′-[γ-S]triphosphate; [γ-S]GTP, guanosine 5′-[γ-S]triphosphate; [γ-S]RNA, RNA containing a γ-thiophosphate at its 5′ end; [γ-S]A-(U-A)$_n$, (A-U)$_n$ initiated with [γ-S]ATP and containing a γ-thiophosphate at its 5′ end; [S]pppXp, nucleoside 5′-[γ-S]triphosphate 2′(3′)-monophosphate; EDTA, ethylenediaminetetraacetate; NaDodSO$_4$, sodium dodecyl sulfate.

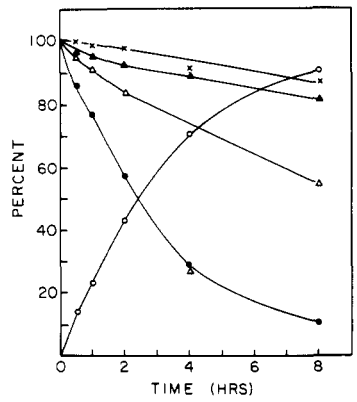

FIGURE 2: Stability of [γ-S]GTP in aqueous solutions. Compounds were incubated in 50 μL of solution at 37 °C. At the indicated time points samples of 5 μL were removed and analyzed by thin-layer chromatography on polyethylenimine as described under Materials and Methods. The breakdown of [8-^{3}H]guanosine 5′-[γ-S]triphosphate was measured in the following solutions: 5 mM Tris-HCl (pH 7.9) (X); 5% acetic acid, 0.5% pyridine (pH 3.5) (△); 7% formic acid (●). The formation of [^{3}H]GDP from radiolabeled [γ-S]GTP was measured in 7% formic acid (○). The breakdown of [^{3}H]GTP was measured in 7% formic acid (▲).

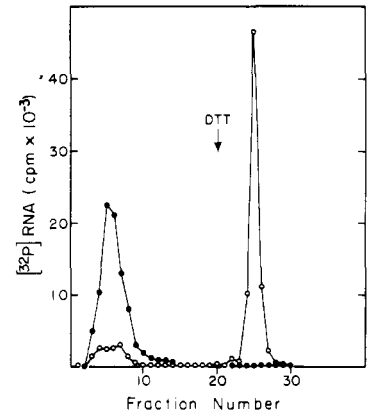

FIGURE 3: Binding of sulfur-containing $(A\text{-}U)_n$ to mercury-agarose. The sulfur-containing RNA [γ-S]A-$(U\text{-}A)_n$, and the control RNA $(A\text{-}U)_n$, were transcribed in 30-s reactions on the template $d(A\text{-}T)_n \cdot d(A\text{-}T)_n$ as described under Methods and Materials. The transcripts were labeled with [α-^{32}P]UTP. The RNA molecules transcribed were chromatographed separately on a mercury-agarose column of freshly regenerated Affi-Gel 501. The addition of 10 mM dithiothreitol to the elution buffer is indicated by the arrow. Radioactivity was determined by counting aliquots of the column fractions in toluene scintillation cocktail containing Triton X-100. (●) $(A\text{-}U)_n$; (○) [γ-S]A-$(U\text{-}A)_n$.

by standard procedures, have a low binding efficiency for RNA larger than approximately 5 S. However, the preparation of special mercury-agarose gels using low concentrations of cyanogen bromide to avoid cross-linking within the gel matrix will allow RNA chains of at least 2000 nucleotides to be retained (Reeve et al., 1977).

Hybridization of [γ-S]RNA with l- and r-Strand λDNA. [γ-S]RNA transcripts were purified by preparative hybridization with separated DNA strands (Bøvre et al., 1971; Blattner & Dahlberg, 1972). Separated strands of λ DNA were prepared as described by Szybalski et al. (1971). Purified [γ-S]RNA synthesized in the presence of [γ-S]ATP was hybridized with a tenfold molar sequence excess of l-strand λ DNA in approximately 0.4 mL of 2 × SSC at 67 °C for 2 h (SSC = 0.15 M NaCl, 0.015 M sodium citrate). Following hybridization the reaction was diluted to 10 mL with 2 × SSC and the hybrids collected on Millipore HAWP nitrocellulose filters by gentle filtration and washed with 2 × SSC. The filtrate was dialyzed into 0.1 M ammonium acetate, lyophilized, and used to hybridize with r-strand DNA. The [γ-S]RNA hybridized with l-strand DNA was melted off the nitrocellulose filter in 2 mL of deionized water at 85 °C for 4 min and lyophilized with 10 μg of yeast carrier RNA in a siliconized test tube.

Chromatography and Electrophoresis. Thin-layer chromatography was performed as described by Goody and Eckstein (1971) on Brinkman polyethylenimine thin layers. Samples were developed in 0.75 M KH_2PO_4 adjusted to pH 3.4 with concentrated HCl. RNase T1 fingerprints, secondary digestions, and sequence analysis were performed by the methods of Sanger (Sanger et al., 1965; Brownlee, 1972). Electrophoresis in the first dimension was performed on cellulose acetate strips in 5% acetic acid, 0.5% pyridine, 5 mM EDTA (pH 3.5), 7 M urea. The second dimension of electrophoresis was done on Whatman DE81 DEAE-paper in 7% formic acid. Autoradiography utilized Kodak BB54 x-ray film.

Results

Stability of Purine Nucleoside 5′-[γ-S]Triphosphates. Initial experiments with [γ-S]ATP and [γ-S]GTP demonstrated some breakdown of the analogues to the corresponding purine nucleoside 5′-diphosphate plus inorganic thiophosphate. Consequently, their stability was studied under a variety of conditions. Both [^{3}H]GTP and tritium radiolabelled [γ-S]-GTP (Materials and Methods) were incubated under various conditions and the conversion to [^{3}H]GDP measured following chromatography on PEI thin layers (Goody and Eckstein, 1971).

Figure 2 illustrates some of these results. [γ-S]ATP and [γ-S]GTP are particularly labile under acid conditions. While 85% of the [^{3}H]GTP is stable after 8 h in 7% formic acid at 37 °C, over 90% of the [γ-S]GTP has been degraded to GDP and thiophosphate. In 5% acetic acid, 0.5% pyridine (pH 3.5), the loss of the γ-thiophosphate is not as rapid, but still over 40% of the compound has been degraded by 8 h. At pH 7.9 in 5 mM Tris-HCl, the nucleoside 5′-[γ-S]triphosphates are relatively stable; after 8 h approximately 10% of the [γ-S]GTP has been converted to GDP. Both sulfur-containing analogues are stable in 0.3 M NaOH for 18 h at 37 °C and store well in neutral aqueous solution frozen at −20 °C. [γ-S]RNA also retains its sulfur when stored as an ethanol precipitate from neutral aqueous solution at −20 °C.

It is clear that prolonged exposure to low pH conditions is to be avoided if retention of the affinity binding property is desired. In particular, loss of the γ-thiophosphate from the initiating 5′-end oligonucleotide is expected during electrophoresis in the second dimension of the two-dimensional fingerprints as this electrophoresis on DEAE paper in 7% formic acid takes from 12 to 18 h. Degradation during the first dimension of electrophoresis in pyridine-acetate (pH 3.5) should be minimal since this step is completed within 60 min (Figure 2).

Binding of [γ-S]RNA to Mercury-Agarose. The ability of mercury-agarose to bind RNA molecules transcribed in a pulse synthesis from initiation complexes was examined using [γ-S]A-$(U\text{-}A)_n$ synthesized from the template $d(A\text{-}T)_n \cdot d(A\text{-}T)_n$. Initiation complexes were formed as described by Blattner & Dahlberg (1972) with *E. coli* RNA polymerase, $d(A\text{-}T)_n \cdot d(A\text{-}T)_n$, and [α-^{32}P]UTP plus either ATP or [γ-S]ATP. Synthesis was started by the addition of Mg^{2+} and terminated after 30 s with EDTA. Rifampin was included in the $MgCl_2$ solution to prevent reinitiation.

Figure 3 illustrates the binding of these transcripts to mercury-agarose. RNA transcribed in the absence of sulfur-

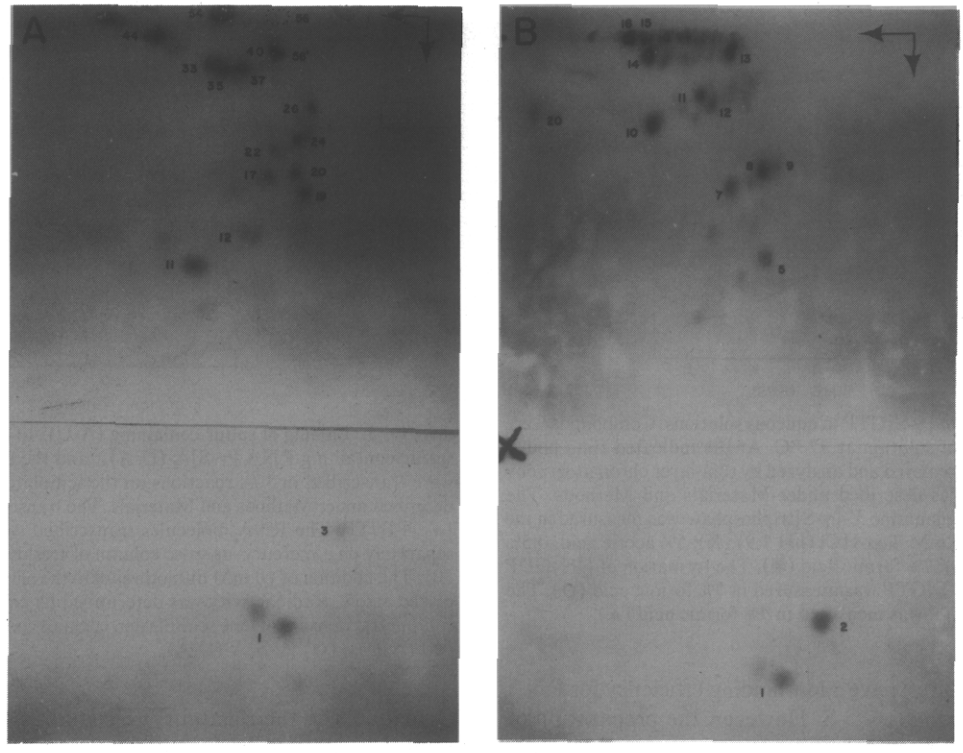

FIGURE 4: (A) RNase T1 oligonucleotide fingerprint of [γ-S]RNA synthesized with [γ-S]ATP, bound to mercury–agarose, and hybridized to l-strand DNA. [γ-S]RNA was transcribed from λcb2 DNA in the presence of [γ-S]ATP using [α-^{32}P]UTP as the labeled nucleotide. It was then purified on Affi-Gel 501 and preparatively hybridized with l-strand λ DNA as described under Materials and Methods. RNase T1 fingerprints were prepared from the hybridized RNA according to standard procedures (Materials and Methods: Brownlee, 1972). The numbering system for the oligonucleotide spots is that presented by Dahlberg & Blattner (1975). The origin is in the upper right hand corner, with the first dimension of migration being right to left and the second dimension being from top to bottom. (B) RNase T1 oligonucleotide fingerprint of [γ-S]RNA synthesized with [γ-S]GTP and isolated on mercury–agarose. [γ-S]RNA was synthesized from λcb2 DNA in the presence of [γ-S]GTP, using [α-^{32}P]UTP as the radiolabeled nucleotide, and isolated on Affi-Gel 501 as described under Materials and Methods. The bound [γ-S]RNA was fingerprinted directly. The origin of the two-dimensional fingerprint is in the upper right hand corner. Electrophoresis in the first dimension is from right to left and the second dimension is from top to bottom. The numbering system for the oligonucleotide spots is arbitrary.

containing nucleotides shows less than 0.1% binding to the column. When the RNA is transcribed using [γ-S]ATP as a precursor, approximately 85% of the incorporated radioactivity is specifically retained on the affinity column. This [γ-S]RNA can be released from the column by the addition of thiol reagents such as 2-mercaptoethanol or dithiothreitol to the elution buffer.

Similar transcriptions were performed using λ DNA as template. [γ-S]RNA was transcribed from λcb2 DNA in a 30-s reaction using either [γ-S]ATP or [γ-S]GTP as one of the four nucleotides and [α-^{32}P]UTP as the radiolabeled nucleotide. The transcripts were then purified over mercury–agarose. Sequences initiated with [γ-S]ATP which hybridize to l-strand λ DNA should specifically represent [γ-S]RNA initiated at the major leftward promoter p_L. Those which hybridize to r-strand DNA should contain a mixture of sequences from both the major and minor rightward promoters p_R and p_R'. [γ-S]RNA initiated with [γ-S]GTP should represent transcripts from the leftward promoter p_O, that is, *oop* RNA (Figure 1).

By these criteria, Table I gives the partition of the synthesized [γ-S]RNA among these promoters. The sum of the [γ-S]GTP- and [γ-S]ATP-initiated RNA bound to the Affi-Gel 501 column is approximately 62% of the total incorporation. Quantitative binding of the [γ-S]RNA to the affinity column was not expected since it has been previously shown that Affi-Gel 501 binds only low molecular weight sequences (Reeve et al., 1977). When the λ transcripts are normalized for this binding efficiency, the distribution of [γ-S]RNA among the various promoters is in good agreement with that determined by hybridization to separated DNA strands of the appropriate λ deletion mutant (Dahlberg & Blattner, 1973).

TABLE I: Distribution of Transcription of [γ-S]RNA by λ Promoters.

Sulfur nucleotide analogue used for initiation	% of total RNA synthesized bound to mercury–agarose	λ promoter	% of total mercury–agarose bound [γ-S]RNA *from* promoter[a]	% *binding* expected[b]
[γ-S]ATP	57	p_L	44	48
		$p_R + p_R'$	37	45
[γ-S]GTP	5.4	p_O	8.7	7

[a] p_L RNA is that [γ-S]RNA initiated with [γ-S]ATP which hybridizes to the separated left strand of λ DNA. Similarly, $p_R + p_R'$ RNA is assumed to be that [γ-S]RNA initiated with [γ-S]ATP which will hybridize with the separated right strand of λ DNA. p_O RNA is that [γ-S]RNA which is initiated with [γ-S]GTP and binds to mercury agarose. [b] Dahlberg & Blattner (1973).

Identification of p_L *[γ-S]RNA.* p_L RNA was analyzed to examine the specific initiation of RNA with [γ-S]ATP. This major leftward RNA codes for the N protein of λ, as well as other genes to the left of N, and is under the control of λ repressor binding to the operator O_L (Ptashne, 1971). The sequence of the first 149 nucleotides of p_L RNA is known (Dahlberg & Blattner, 1975) as well as 85 base pairs of DNA

TABLE II: RNase T1 Oligonucleotides of Left Strand [γ-S]RNA Initiated with [γ-S]ATP.[a]

Oligo-nucleotide no.	RNase A digestion products	Presumed structure	Rel molar yield
1	Gp*	Gp*[U]	1.2
3	Gp*	C-C-Gp*[U]	0.5
11	Gp*	U-Gp*[U]	1.5
12	Cp*	C*-U-Gp	0.4
17	A-A-Gp*	U-A-A-Gp*[U]	0.5
19	Cp*	C-C-C*-U-Gp	1.0
20	A-Cp*	C-A-C*-U-Gp	1.0
22	A*-Up	A*-U-A-C-Gp	0.3
24	A-A*-Up, Gp* (1:1)	C-C-A-A*-U-Gp*[U]	0.5
26	A*-Up	A-C-C-A-C-C-A*-U-Gp	0.9
33	A*-Up*	A*-U*-U-A-Gp	0.5
35	A*-Up*	C-A*-U*-U-Gp	0.9
37	A*-Up*	A*-U*-U-C-C-Gp	0.7
40 + 56′	A*-Up*, ppX*-Up (2:1)	C-A*-U*-U-C-A-A-A-Gp, ppA*-U-C-A-Gp	0.9
44	Up*, Cp* (2:1)	C*-U*-U*-U-Gp	0.7
54	Cp*, Up*, A-A-A-A-A*-Up (2:1:2)	C*-U-C*-U*-U-A-A-A-A-A*-U*-U-A-A-Gp	1.0

[a] The numbered oligonucleotides of Figure 4A were cut from the DEAE-paper using the autoradiogram as a template. The radioactivity of the spots was determined by counting Cerenkov radiation in a scintillation counter. The oligonucleotides were eluted with 2 M ammonium acetate containing carrier yeast oligonucleotides, lyophilized, and subjected to secondary digestions by standard procedures with either RNase A or 0.3 M NaOH (Brownlee, 1972). The relative molar yields of the oligonucleotides were calculated using the radioactivity of spot 19 and of spot 20 as being 1.0 each, and assuming the structures of the spots to be as indicated. The radioactive phosphorus atom is designated by an asterisk (*) and brackets [] denote the 3′ nearest neighbor nucleotide.

in the O_L operator region preceding the 5′-end of p_L RNA (Maniatis et al., 1974; Dahlberg & Blattner, 1975; Ptashne et al., 1976).

Transcripts from λcb2 DNA were synthesized in the presence of [γ-S]ATP, with [α-^{32}P]UTP as radiolabel, and isolated on mercury–agarose. The retained [γ-S]RNA was then hybridized with separated l-strand λ DNA in solution and collected on nitrocellulose filters as described by Bøvre et al. (1971). The [γ-S]RNA was melted off the filters and subjected to RNase T1 fingerprint analysis (Sanger et al., 1965; Brownlee, 1972). Figure 4A shows the observed pattern of [α-^{32}P]UTP-labeled T1 oligonucleotides.

The overall features of this fingerprint are in agreement with the sequence and fingerprint data for p_L RNA determined by Dahlberg & Blattner (1975). Table II presents the results of the secondary digestion of the T1 oligonucleotides with RNase A. The positions of the labeled phosphates were further assigned from the products resulting from complete alkaline digestion of the T1 spots. The observed digestion products are in complete agreement with expected sequences of p_L RNA.

Two differences from the standard fingerprint (Dahlberg & Blattner, 1975) are observed. There is an additional unnumbered oligonucleotide to the left of spot 44. The source of this spot is not known but it might arise from transcription past nucleotide 149 since our reactions were for 30 s instead of 20 s as used by Blattner & Dahlberg (1972). The second difference in the fingerprint is the absence of the initiating 5′-end oligonucleotide [S]pppApUpCpApGp from its position as spot 56. Altered mobility of the 5′-terminal oligonucleotide is expected as a result of the instability of the γ-thiophosphate, particularly in the second dimension of the fingerprint where electrophoresis was done in 7% formic acid (Figure 2). Loss of the thiophosphate and its associated charges should result in a greater electrophoretic mobility of spot 56 in the second dimension on DEAE paper. Apparently it comigrated with spot 40. Alkaline hydrolysis of spot 40 resulted in the release of labeled ppXp and RNase A digestion produced a highly charged product presumed to be ppXpUp in addition to the expected products from the sequence of spot 40. We have designated this oligonucleotide containing the 5′-diphosphate as spot 56′ (Table II).

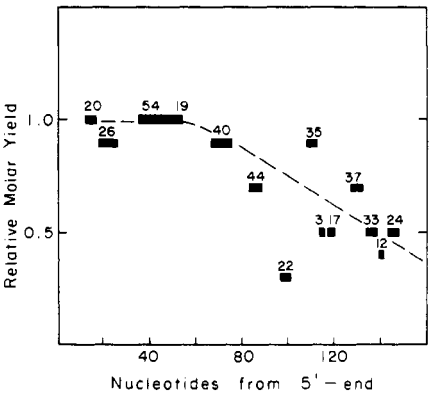

FIGURE 5: Relative molar yields of p_L [γ-S]RNA T1 oligonucleotides as a function of their distance from the 5′ end. The relative molar yields (Table II) of the T1 oligonucleotides which occur uniquely in the sequence of p_L RNA have been plotted as a function of their position from the 5′ end of the sequence (Dahlberg & Blattner, 1975).

The unique T1 oligonucleotides from p_L RNA have been ordered by their distance along the RNA sequence and the observed molar yields have been plotted against their position in the sequence (Figure 5). The polarity of the molar yields is consistent with initiation of p_L RNA at, or near, its authentic startpoint.

We conclude that [γ-S]RNA transcribed in the presence of [γ-S]ATP, isolated on mercury–agarose, and hybridized to l-strand γDNA, represents sequences of p_L RNA. The transcript initiated at, or near, the correct nucleotide with the probable sequence [S]pppApU. Furthermore, no major errors in the fidelity of transcription occurred when ATP was replaced by [γ-S]ATP.

Identification of oop [γ-S]RNA. The only RNA known to initiate with GTP in vitro from λcb2 DNA is the 4S RNA transcribed from the minor leftward promoter p_O at about 80% on the λ physical map. This 4S RNA, designated *oop* RNA (Hayes & Szybalski, 1973), is a self-terminating RNA 77 nucleotides long. Its sequence is known (Dahlberg & Blattner, 1973; Scherer et al., 1977). In addition, the sequence of 59 base pairs of DNA preceding the 5′ end of the RNA (Scherer et al., 1977) and 35 nucleotides beyond the 3′ end of the transcript (Rosenberg et al., 1976) have also been determined.

Since *oop* RNA is the only transcript initiated with GTP and represents a minor fraction of the total incorporation, it provides an excellent test of the specificity of the affinity binding system. [γ-S]RNA transcribed from λcb2 DNA in the presence of [γ-S]GTP, with [α-^{32}P]UTP as the radiolabeled nucleotide, was isolated on mercury–agarose as before. The retained [γ-S]RNA was examined by fingerprint analysis of its T1 oligonucleotides (Figure 4B). Table III presents the catalogue of RNase A and alkaline digestion products derived from each T1 spot.

The positions of the numbered spots correspond exactly to

TABLE III: RNase T1 Oligonucleotides of *oop* [γ-S]RNA.[a]

Oligonucleotide no.	RNase A digestion products	Presumed structure	Rel molar yield
1	Gp*	Gp*[U]	0.9
2	Gp*	C-Gp*[U]	0.8
5	Cp*	C*-U-C-Gp	0.7
7	A*-Up	A*-U-A-Gp	0.9
8	A*-Up, A-Gp* (1:1)	A*-U-C-C-A-Gp*[U]	0.6
9	Cp*	A-C-C*-U-C-A-Gp	0.9
10	Up*	U*-U-Gp	1.1
11	A-A*-Up	U-A-A*-U-Gp	1.1
12	Up*	U*-U-C-A-Gp	0.9
13	Cp*, A*-Up, A-A-Cp* (1:1:1)	A-A-C*-U-C-C-A*-U-C*-U-Gp	0.9
14	Up*, Gp*, A*-Up* (1:1:2)	A*-U*-U*-U-Gp*[U]	1.0
15	Up*	U*-U*-U*-U*-U*-U-A_{OH}	0.3
16	Up*, A*-U_{OH} (5:1)	U*-U*-U*-U*-U*-U-A*-U_{OH}	0.6
20	ppGp*	ppGp*[U]	0.5

[a] The numbered oligonucleotides of Figure 4B were cut from the DEAE-paper electropherogram and analyzed as for Table II. The relative molar yields were calculated using the radioactivities of spot 11 and spot 12 to be 1.0 each, and assumed the structures of the spots to be as indicated.

TABLE IV: Degradation Products of the Oligonucleotide Bound to Mercury-Agarose following RNase A Digestion of *oop* [γ-S]RNA.[a]

Treatment of bound oligonucleotide	Degradation products	Fraction of total radioactivity
Alkaline digestion	Up*	0.43
	[S]pppXp*	0.38
	Gp*	0.19
Sequential digestion with HCOOH, alkaline phosphatase, and RNase T1	[S]pppXp*	0.1
	Gp*	0.4

[a] *Oop* [γ-S]RNA isolated on mercury-agarose was digested with RNase A and the sulfur-containing oligonucleotide was purified by a second chromatography on mercury-agarose. It was eluted in 0.1 M ammonium acetate, 10 mM 2-mercaptoethanol and lyophilized with carrier yeast oligonucleotides. Alkaline hydrolysis was performed in 0.3 M NaOH, 37 °C, for 18 h. Sequential digestion involved hydrolysis with 7% formic acid, 37 °C, for 6 h. The sample was lyophilized and digested with alkaline phosphatase at 1 part enzyme to 5 parts nucleotide for 60 min at 37 °C in 10 mM Tris-HCl (pH 8). Phosphatase was inactivated by heating at 90 °C for 3 min with 1 mM EDTA. Finally, the sample was digested with RNase T1 for 30 min at 37 °C. Products were analyzed by electrophoresis in pyridine acetate (pH 3.5) on Whatman 3MM and DE81 papers.

oop RNA oligonucleotide sequences expected to be labeled by [α-^{32}P]UTP. The secondary digestion products are also in agreement with the assignments of the T1 spots to *oop* RNA sequences.

Oligonucleotide spots 15 and 16 contain 3′-OH termini as determined by the rate of partial digestion by snake venom phosphodiesterase. Spot 16 contains the terminal oligonucleotide $(U)_6pApU(OH)_2$ and accounts for 75% of the terminated transcripts. A similar termination has been seen by others (Rosenberg et al., 1975; Howard et al., 1977). Spot 15 is presumed to be the sequence $(U)_6pA(OH)_2$ and accounts for about 25% of the transcripts. This was the major terminal oligonucleotide obtained by Dahlberg and Blattner (1973).

Spot 20 is the nucleoside triphosphate ppGp, the presumptive initiating nucleotide. The absence of the γ-thiophosphate is in accordance with the instability of the analogue in formic acid (Figure 2), and is similar to the results obtained with the initiating oligonucleotide of p_L RNA. Since this nucleotide contains a radioactive 2′(3′)-phosphate, the [γ-S]RNA must have initiated with the sequence [S]pppGpU.

The exact initiation sequence of the [γ-S]RNA initiated with [γ-S]GTP was further studied under conditions where the γ-thiophosphate would not be lost. Transcripts synthesized in the presence of [γ-S]GTP were purified by binding to mercury-agarose. The retained [γ-S]RNA was eluted from the column and subjected to RNase A digestion. This digest was then again applied to the mercury-agarose column to bind the sulfur-containing RNase A oligonucleotides. Of the total radioactivity in the digest, 7.2–7.6% was retained on the affinity column which agrees closely with the predicted binding. The *oop* RNA sequence contains 24–25 radiolabeled UMP residues, depending on termination, and the RNase A 5′-end oligonucleotide [S]pppGp*Up* contains two of these radiolabeled phosphates (Scherer et al., 1977). Therefore about 8.1% of the radioactivity in the RNase A digest is expected to bind to the mercury-agarose column.

The RNase A oligonucleotide bound to the mercury-agarose affinity column was analyzed by further digestion (Table IV). Complete alkaline digestion released a nucleoside tetraphosphate [S]pppXp and Up in about equal molar yields. A smaller amount of Gp was also obtained. From this nearest neighbor transfer analysis, it follows that the initiating nucleotide is followed by a uridine residue which in turn must be the 3′-terminal nucleotide of the RNase A oligonucleotide. The release of an equal amount of Up indicates that this uridine was itself followed by a uridine residue in the RNA. Therefore, the [γ-S]RNA must have initiated with the sequence [S]-pppXpUpU.

Characterization of the nucleoside tetraphosphate is complicated by the fact that the γ-thiophosphate makes the 5′-phosphates resistant to bacterial alkaline phosphatase attack (Goody & Eckstein, 1971). Therefore, the γ-thiophosphate was first removed by treatment with 7% formic acid at 37 °C for 6 h. The unprotected oligonucleotide was then sequentially digested with alkaline phosphatase and RNase T1. The products of this sequential digestion were a major spot of Gp and a trace of nucleoside tetraphosphate probably resulting from incomplete removal of the sulfur by formic acid (Figure 2). Approximately one-half of the radioactivity was not recovered as a nucleotide product and was presumed to be released as inorganic phosphate. The fact that the only radioactive phosphate protected from alkaline phosphatase digestion was recovered as Gp identifies the initiating nucleotide as [γ-S]GTP. These results are consistent with the initial sequence of the [γ-S]RNA being [S]pppGpUpU, the authentic initiating sequence of *oop* RNA.

We conclude that mercury-agarose specifically retains RNA containing *oop* RNA sequences when λcb2 DNA is transcribed in the presenle of [γ-S]GTP. Furthermore this affinity bound [γ-S]RNA initiated with the probable sequence [S]pppGpUpU. As with p_L RNA no major errors in the fidelity of transcription occurred as a result of the replacement of GTP with [γ-S]GTP.

Discussion

The characterization of the λ [γ-S]RNA transcripts presented here has relied upon the oligonucleotides labeled with [α-^{32}P]UTP. Consequently, the sequences of the isolated transcripts are not unambiguously defined. However, both the RNase T1 fingerprint patterns and the products derived from the secondary digestions of the T1 spots with RNase A and alkali correspond exactly to the known fingerprints and expected digestion products of p_L RNA and *oop* RNA. There is therefore a very high probability that the assignments of these transcripts to p_L RNA and *oop* RNA are correct.

The fact that the general distribution of [γ-S]RNA sequences among the various promoters corresponds to the distribution seen with nonsulfur RNA suggests that substitution of the thiol analogues does not greatly change the relative efficiencies of the λ promoters. The correspondence of the T1 oligonucleotides and their digestion products to known sequences indicates that there is no gross infidelity of transcription caused by the substitution of the sulfur-containing nucleotides. There can also be no major reduction in the rate of elongation for in the case of p_L RNA the fingerprint shows the presence of nucleotides out to position 149 which is the length obtained by Blattner & Dahlberg (1972) for their transcriptions. Furthermore, the results suggest that the transcripts synthesized in the presence of the thiol analogues initiated at their correct startpoints.

The analysis of the major leftward [γ-S]RNA indicates that initiation occurred near the correct 5′ end with the sequence [S]pppApU. In addition to the authentic initiation site, this sequence also occurs 11 nucleotides preceding the correct startpoint (Dahlberg & Blattner, 1975; Maniatis et al., 1974; Ptashne et al., 1976). A T1 digest of RNA with this incorrect initiation would give the 5′-end oligonucleotide ppApUpApGpUpGp and subsequent digestion with RNase A would yield the labeled dinucleotide ApCp. In fact, this dinucleotide was not obtained from the analysis of spots 40 plus 56′ and it therefore seems likely that the bulk of p_L [γ-S]RNA initiated at its correct startpoint.

In the case of *oop* [γ-S]RNA, the initial sequence was determined to be [S]pppGpUpU. This sequence of three nucleotides is relatively rare. Aside from the correct initiation site this sequence does not occur until 37 nucleotides preceding the correct startpoint (Scherer et al., 1977). Initiation at this site would contribute 18 additional radiolabeled phosphates to the fingerprint which are not obtained. Furthermore, the transcripts initiated at this site at −37 bases would begin [S]-pppGpUpUpGpApGpU and continue. An RNase T1 digest of such RNA would give a high yied of the labeled dinucleotide ApGp and this spot is absent from the fingerprints. We therefore believe that the bulk of the *oop* RNA was initiated at its correct site.

It is still possible that sequences were initiated incorrectly with the sulfur analogues but that these initial sequences were not labeled by [α-^{32}P]UTP and could not be detected. This is unlikely, however, because of the relative molar yields of the identified initiating oligonucleotides. Good yields were obtained especially considering the diffusion of the spots due to the progressive loss of the γ-thiophosphate during electrophoresis in 7% formic acid. For example, Dahlberg & Blattner (1975) report a yield of 0.5 for the 5′-end oligonucleotide of p_L RNA. In the present case, estimation of the yield of spot 56′ is complicated by its comigration with spot 40. Nevertheless, the yield of the radioactive phosphate in spot 56′ is at least as good as either of the two present in spot 40 (Table II). For *oop* RNA a yield of 0.5 was obtained for the ppGp of spot 20. Moreover, the 5′-end oligonucleotide produced by RNase A digestion was recovered from the mercury–agarose column in better than 0.9 molar yield. These considerations suggest that the 5′-end sequences analyzed represent the major modes of initiation of the [γ-S]RNA molecules isolated.

While the fingerprint of p_L RNA is very clean, there are some additional faint spots in *oop* RNA. Approximately 80% of the radioactivity of the fingerprint can be accounted for by the numbered T1 oligonucleotides and belong to the sequence of 77 nucleotides of *oop* RNA. The remaining radioactivity belongs to distinct spots obtained in molar yields of 0.2 and less. There are several possible sources of these additional spots. The majority of them undoubtedly arise from read-through of the termination point of *oop* RNA. Howard et al. (1977) have recently measured the efficiency of termination of *oop* RNA and report that, in the absences of the *E. coli* protein *rho*, about 14–15% of the transcripts fail to terminate at nucleotides 77 or 78. If these read-through transcripts resulted in RNA of the same size as p_L RNA (140–150 nucleotides), the additional oligonucleotides would comprise about 10–15% of the radioactivity of the fingerprint. Other sources of extra faint spots include initiation with [γ-S]GTP at incorrect sites by either holoenzyme or traces of core polymerase (Bautz et al., 1969; Johnston & McClure, 1976) and specific initiation with [γ-S]GTP at weak promoter sites as yet uncharacterized. The latter possibility is particularly intriguing and candidates for such sites might be the promoter p_I in the gene *xis* (Shimada and Campbell, 1974a,b), or the promoter in gene *rex* responsible for *lit* RNA transcription (Hayes & Szybalski, 1973). Another possible source of the extra spots is the species 3 RNA transcribed from the l-strand of λ (Smith & Hedgpeth, 1975).

Since thiol nucleotides can initiate accurately at defined promoter sites, they represent extremely useful affinity probes for the study of RNA chain initiation. The ability to selectively bind sequences initiated in vitro results in amplification of the initiation event. For example, the incorporation of [γ-^{32}P]GTP into *oop* RNA would yield a single radioactive phosphate per chain initiation. In contrast, synthesis with [α-^{32}P]UTP plus [γ-S]GTP, and subsequent affinity chromatography, results in 24–25 radioactive phosphates detected per initiation event. Longer chain length RNA molecules and additional radiolabeled nucleotide precursors would result in a correspondingly greater sensitivity of detection.

A potential problem when applying these analogues to transcription in isolated nuclei or chromatin is that the γ-thiophosphate might be transferred to other nucleoside or nucleotide precursors. While this must be examined for each synthetic system, it does not occur in isolated mouse plasmacytoma tissue culture nuclei. We have recently shown that ribosomal 5S RNA transcribed in these nuclei in vitro can be bound to mercury–agarose only when [γ-S]GTP is used as a substrate and not when [γ-S]ATP is the sulfur nucleotide (Smith et al., in preparation). Therefore the sulfur label could not have been transferred to any significant extent from [γ-S]ATP to either the α-phosphate position of any of the nucleoside triphosphates or to the β- or γ-phosphate position of GTP.

The ability to physically isolate in vitro primary transcripts should be particularly useful. It will allow separation of RNA actually initiated in vitro from that which results merely from the extension of preexisting nascent chains (Smith & Huang, 1976; Shih et al., 1977). By determining the analogue dependent binding of an RNA sequence, the exact nucleotide used to initiate that sequence in vitro can be deduced directly. Finally, it should complement in vivo studies of mRNA precursor

molecules and define these precursors in vitro in an unambiguous manner.

Acknowledgments

We wish to thank R. P. Dottin, K. Murray, F. R. Blattner, P. C. Huang, and M. Rhoades for helpful discussions during the course of this work.

References

Bautz, E., Bautz, F., & Dunn, J. (1969) *Nature (London) 223,* 1022–1024.

Blattner, F. R., & Dahlberg, J. E. (1972) *Nature (London), New Biol. 237,* 227–232.

Bøvre, K., Lozeron, H. A., & Szybalski, W. (1971) in *Methods in Virology* (Moramorosch, K., & Koprowski, H., Eds.) Vol. 5, pp 271–292, Academic Press, New York, N.Y.

Brownlee, G. G. (1972) *Determination of Sequences in RNA,* North-Holland Publishing Co., Amsterdam.

Chamberlin, M. J. (1974a) *Enzymes, 3rd Ed. 10,* 333–374.

Chamberlin, M. J. (1974b) *Annu. Rev. Biochem. 43,* 721–775.

Dahlberg, J. E., & Blattner, F. R. (1973) in *Virus Research* (Fox, C. F., & Robinson, W. S., Eds.) pp 533–543, Academic Press, New York, N.Y.

Dahlberg, J. E., & Blattner, F. R. (1975) *Nucleic Acids Res. 2,* 1441–1458.

Goody, R. S., & Eckstein, F. (1971) *J. Am. Chem. Soc. 93,* 6252–6257.

Goody, R. S., Eckstein, F., & Schirmer, R. H. (1972) *Biochim. Biophys. Acta 276,* 155–161.

Hatlen, L. E., Amaldi, F., & Attardi, G. (1969) *Biochemistry 8,* 4989–5005.

Hayes, S., & Szybalski, W. (1973) *Mol. Gen. Genet. 126,* 275–290.

Howard, B. H., deCrombrugghe, B., & Rosenberg, M. (1977) *Nucleic Acids Res. 4,* 827–842.

Johnston, D. E., & McClure, W. R. (1976) in *RNA Polymerase* (Losick, R., & Chamberlin, M., Eds.) pp 413–428, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Maitra, U., & Hurwitz, J. (1965) *Proc. Natl. Acad. Sci. U.S.A. 54,* 815–822.

Maniatis, T., Ptashne, M., Barrell, B. G., & Donelson, J. (1974) *Nature (London) 250,* 394–397.

Marzluff, W. F., Jr., Murphy, E. C., Jr., & Huang, R. C. C. (1974) *Biochemistry 13,* 3689–3696.

Ptashne, M. (1971) in *The Bacteriophage Lambda* (Hershey, A. D., Ed.) pp 221–237, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Ptashne, M., Backman, K., Humayun, M. Z., Jeffrey, A., Maurer, R., Meyer, B., & Sauer, R. T. (1976) *Science 194,* 156–161.

Reeve, A. E., Smith, M. M., Pigiet, V., & Huang, R. C. C. (1977) *Biochemistry 16,* 4464–4469.

Roberts, J. (1969) *Nature (London) 224,* 1168–1174.

Rosenberg, M., Weissman, S., & deCrombrugghe, B. (1975) *J. Biol. Chem. 250,* 4755–4764.

Rosenberg, M., deCrombrugghe, B., & Musso, R. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73,* 717–721.

Sanger, F., Brownlee, G. G., & Barrell, B. G. (1965) *J. Mol. Biol. 13,* 373–398.

Scherer, G., Hobom, G., & Kossel, H. (1977) *Nature (London) 265,* 117–121.

Shih, T. Y., Young, H. A., Parks, W. P., & Scolnick, E. M. (1977) *Biochemistry 16,* 1795–1801.

Shimada, K., & Campbell, A. (1974a) *Proc. Natl. Acad. Sci. U.S.A. 71,* 237–241.

Shimada, K., & Campbell, A. (1974b) *Virology 60,* 157–165.

Smith, G. R., & Hedgpeth, J. (1975) *J. Biol. Chem. 250,* 4818–4821.

Smith, M. M., & Huang, R. C. C. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73,* 775–779.

Szybalski, W., Kubinski, H., Hradecna, Z., & Summers, W. C. (1971) *Methods Enzymol. 21,* 383–413.