Sequence of the promoter-operator proximal region of the major leftward RNA of bacteriophage lambda

James E. Dahlberg* and Frederick R. Blattner[#]

*Department of Physiological Chemistry, University of Wisconsin, Madison, WI 53706, USA

Received 23 April 1975

ABSTRACT

The sequence of the first 149 nucleotides of the major leftward RNA of bacteriophage lambda has been determined. Preliminary sequence information was also obtained for a portion of the untranscribed area immediately upstream of the point on the template when RNA synthesis normally starts. Several restriction endonuclease sites, deletion endpoints, and single base changes have been localized within the sequence. The first potential translation initiation codon which is not followed by an in-phase termination codon is a GUG located 90 nucleotides from the transcription startpoint.

INTRODUCTION

Regulation of gene expression in bacteria and bacteriophage depends on specific sequences in DNA which are recognized by proteins (1-3). Transcription of the major leftward operon of bacteriophage λ is regulated by a promoter-operator region, p_L - o_L (4-6). This region contains DNA sequences specific for binding and initiation by RNA polymerase as well as sequences for binding regulatory proteins such as the λ repressor and presumably, the *tof* (*cro*) gene product. We have previously shown that the promoter and operator mutation sites are located upstream of the point where RNA starts, s_L , and are not transcribed into RNA (6). The size of this controlling area, as measured by repressor binding, is about 115 nucleotide pairs and it contains six repressor binding sites in tandem (7). The promoter mutant *sex1* and operator mutant *y2* are located about 35 base pairs from s_L according to recent measurements (8-11), although our earlier measurement had erroneously placed them further upstream (12).

In this communication we present the results of nucleotide sequence analysis of this area. In the region downstream from s_L we have determined the complete sequence of the first 149 nucleotides. In the upstream area we used transcription of the complimentary strand to obtain a catalogue of oligonucleotides for a portion of the untranscribed control region. Our analysis of the untranscribed region is in agreement with the sequence determined by

Maniatis et al. (4), and allows us to determine the probable base change resulting from the y2 operator mutation.

MATERIALS AND METHODS

DNA Preparations.

DNAs of bacteriophage λcb2 and λbio30-7dvlnin5 were prepared by phenol extraction of the phage. Supercoiled λdv DNA was isolated from E. coli carrier cells by the lysis procedure of Vapnek and Rupp (13), followed by centrifugation in a CsCl gradient containing ethidium bromide. Linear λdvl DNA was prepared by cutting λbio30-7dvlnin5 DNA (6) with Eco R1. Since this enzyme cuts λ dvl in only one place, the duplication phage gives a fragment that is indistinguishable from Eco R1 digested λdvl circles.

DNA fragments resulting from restriction endonuclease digestion were separated by electrophoresis through 1% agarose slab gels (in electrophoresis buffer consisting of 40 mM Tris-acetate pH 8, 20 mM Na acetate, 2 mM Na₂EDTA, 18 mM NaCl) in an EC 470 electrophoresis cell (E.C. Corp., St. Petersburg, Fla.) using conditions developed by R. Helling et al. (14). DNA bands were visualized by staining the gel 15-60 minutes in 0.5 µg/ml ethidium bromide. The bands were cut out and DNA was extracted by soaking macerated gel slices in electrophoresis buffer. The preparation was then frozen and thawed, centrifuged to remove gel pieces, filtered through a Millipore filter (HAWP-0.45µ) and concentrated by precipitation with four volumes of ethanol at -20°C. The DNA was resuspended in 5 ml buffer, (0.1 M Tris pH 8, 0.05 M NaCl) and incubated for 10 min at 37°C with 25 microliters of diethyl-pyrocabonate, to insure against RNase contamination. After evacuation to remove evolved CO₂, DNA was again precipitated with 4 volumes of ethanol and resuspended in 0.05 ml of buffer (0.01 M Tris pH 8, 0.01 M NaCl).

Enzymes, Carrier RNA, and Triphosphates.

E. coli RNA polymerase was prepared by the procedure of Chamberlin and Berg and then centrifuged through a low salt glycerol gradient (16). RNase T₁ and pancreatic RNase were purchased from Calbiochem and Worthington Biochemicals, respectively. Restriction endonucleases were gifts from B. Allet, A. Landy, D. Nathans, L. Robinson, and B. Weisblum. Carrier RNA was purchased from Sigma (Type VI yeast RNA) and was extracted with phenol before use.

Unlabeled ribonucleoside-triphosphates ATP, GTP, CTP and UTP were purchased from Schwarz Biochemicals, Inc. α³²P-labeled nucleoside-triphosphates were prepared by our modification of the method of Symonds (17).

Preparation of RNAs for Sequencing.

Two types of in vitro RNA transcripts were prepared which we call "sense" and "antisense" transcripts. These designations refer to the direction of in vitro transcription relative to what normally occurs in vivo.

Sense transcripts were synthesized in vitro using intact λ cb2 DNA as template. The standard preincubation mixture contained 200 μ g of RNA polymerase, 200 μ g of DNA, 50 nmol of each nucleoside triphosphate (one labeled in the α position with 1-10 mCi of 32 P), 4 μ mol of EDTA, 2 μ mol of $MgCl_2$, 25 μ mol of Tris, pH 7.9, and 85 μ mol of KCl, all in a volume of 0.5 ml. After 2 min at 25°C the reaction was started by adding 20 μ mol of $MgCl_2$ mixed with 5 μ g of rifampin in 0.05 ml of H_2O . The reaction was stopped by adding 50 μ mol of EDTA. The RNA chain length was kept short (about 150 nucleotides) by limiting the period of synthesis to 0.4 minutes. In such products, the 5' end corresponded to the normal in vivo RNA startpoint (18, 19), and the 3' end was determined by the duration of synthesis. Radioactive RNA was purified by sephadex G50 chromatography, DNase digestion, diethylpyrocarbonate treatment, ethanol precipitation, and sequential steps of preparative hybridization with elution in water at 85°C as described in Figs. 2 and 3 of reference 6. RNA prepared in this way provided most of the information needed to determine the sequence immediately downstream from the startpoint s_L.

Antisense RNA was synthesized on the λ dv DNA HinII restriction fragment containing the novel joint which is described below. The conditions for transcription of the fragment were the same as for sense RNA except that no rifampin was used, and the time of incubation was 2 hours. Since the fragment was only 3% as long as intact λ DNA, correspondingly less DNA was added per reaction mixture.

Antisense RNA was purified in the same manner as sense RNA except that only one step of preparative hybridization to r strands of λ nin5 was used and the hybrids were digested on the filters with pancreatic RNase to remove RNA corresponding to the nin area of the template. After treatment of the filters with iodoacetate to inactivate RNase, the RNA was eluted with 5 ml of 95°C H_2O , 100 μ g of carrier RNA was added and the solution was incubated with 25 μ l of diethylpyrocarbonate at 37°C for 10' to completely inactivate any RNase. RNA then was precipitated with ethanol and sequenced.

Sequencing of the RNA.

Determination of the nucleotide sequence of the RNA was done by standard nearest neighbor phosphate transfer sequencing techniques (21,22). The RNA

was digested with either RNase T₁ or pancreatic RNase and the resulting oligonucleotides were separated by two-dimensional high voltage paper electrophoresis (23). Conditions of partial digestion were either 0.1 µg RNase T₁ and 100 µg RNA (primarily carrier RNA) incubated in 3 µl of 0.01 M Tris HCl, pH 7.6, 0.01 M MgCl₂ for 6 minutes at 22°C or 2 µg carboxymethylated pancreatic RNase (25) and 60 µg RNA in 2 µl 0.05 M Na acetate, pH 4, 0.01 M EDTA for 45 minutes at 37°C. Partial nuclease digestion products were purified by electrophoresis on cellulose acetate at pH 3.5 and homochromatography (24). Products were eluted and analyzed by T₁ or pancreatic RNase digestion and each oligonucleotide was characterized by electrophoretic mobility on DEAE cellulose paper (in 7% formic acid) and by redigestion with pancreatic or T₁ RNase, as appropriate. In some cases, alkaline digestion of the redigestion products was used to ascertain the exact position of the radioactive phosphate.

RESULTS

Nucleotide Sequence of Sense RNA.

Oligonucleotides produced by digestion of the major leftward RNA with RNase T₁ or pancreatic RNase were separated by two-dimensional paper electrophoresis (fingerprinting). Representative fingerprints have been published elsewhere (6). Figures 1a and b are schematic representations of the patterns, showing the numbering systems that were used.

Tables 1 and 2 present catalogues of RNase T₁ and pancreatic RNase oligonucleotides that were obtained from the first 149 nucleotides of the major leftward RNA of bacteriophage λ. The tables show the nearest neighbor phosphate labeling patterns of each redigestion product of each oligonucleotide. In general, the sequence of each oligonucleotide was deduced from these analyses. One T₁ RNase oligonucleotide, number 54, presented a special problem because of its length and high content of pyrimidines. The nucleotide sequence of that oligonucleotide was facilitated by analysis of oligonucleotides A46 and A117 of antisense RNA (see below).

The entire sequence was reconstructed using nearest neighbor phosphate transfer analysis and partial RNase digestions for overlapping of T₁ and pancreatic RNase products. Figure 2 shows the sequence, as well as partial digestion products which were useful in ordering the oligonucleotides. In addition, RNA products which were only about 30-45 nucleotides long were produced by shortening the time of synthesis. These products were useful in confirming the relative order of oligonucleotides near the 5' end. For

Figure 1. Fingerprints obtained from RNase T₁ (a and c) or pancreatic RNase (b and d) digestion of *in vitro* synthesized sense (a and b) and antisense (c and d) RNA. Figure 1a and b are schematic representations of fingerprints of major leftward RNA the originals of which were published previously (6). Figure 1c and d are photographs of fingerprints of $\alpha^{32}\text{P}$ -GTP labeled antisense RNA. The arrows in d indicate positions of oligonucleotides 102 and 106B which are not labeled by GTP. The numbering of spots corresponds to that used in the tables except that the prefix "A" was omitted from c and d for the sake of clarity. The position of the blue dye marker (xylene cyanol-FF) is shown by a square. Pancreatic RNase products UMP (109 and A109) are not shown because they migrated too far in the second dimension. The first dimension of electrophoresis (pH 3.5) is from right to left, and the second (7% formic acid) is from top to bottom.

Table 1

RNase T₁ OLIGONUCLEOTIDES DERIVED FROM THE MAJOR LEFTWARD (SENSE) RNA

Oligo-nucleotide number	GTP	ATP	UTP	CTP	Relative Molar Yield	Structure
1	3.2 [*] Gp	2.1 [*] Gp	4.7 [*] Gp	3.3 [*] Gp	3.2	Gp[G]
					2.1	Gp[A]
					4.7	Gp[U]
					3.3	Gp[C]
2	0.7 [*] Cp	0.5 [*] Gp			0.6	C-Gp[A]
3	1.0 [*] Cp	----	1.3 [*] Gp	1.4 [*] Cp	1.3	C-C-Gp[U]
4	4.1 [*] A-Gp (3:1)	4.6 [*] Cp, A-Gp (3:1)	----	1.0 A-Gp [*]	1	C-A-Gp[G]
					1.2	C-A-Gp[A]
					1.0	C-A-Gp[C]
5	2.1 A-Gp [*]	----	----	2.0 [*] Gp, ApCp	2.0	A-C-Gp[C]
7	8.0 A-A-Gp [*] (3:2)	5.9 A-A-Gp [*] (5:1)	----	0.7 A-A-Gp [*]	3	A-A-Gp[G]
					1	A-A-Gp[A]
					0.7	A-A-Gp[C]
8	0.7 A-A-A-Gp [*]	1.8 A-A-A-C-Gp [*] (2:1)	----	0.6 A-A-A-Gp [*]	0.7	A-A-A-C-Gp[A]
11	4.6 Up [*]	1.9 [*] Gp	1.9 [*] Gp	0.9 [*] Gp	1.9	U-Gp[A]
					1.9	U-Gp[U]
					0.9	U-Gp[C]
12	0.4 Up [*]	----	0.5 [*] Cp	0.4 [*] Gp	0.4	C-U-Gp[C]
17	0.7 A-A-Gp [*]	1.8 A-A-Gp, Up [*]	1.0 A-A-Gp [*]	----	0.9	U-A-A-Gp[U]
19	1.0 Up [*]	1.1 [*] Gp	1.0 [*] Cp	2.0 [*] Cp	1.0	C-C-C-U-Gp[A]
20	1.0 Up [*]	2.3 [*] Gp, [*] Cp	1.1 A-Gp [*]	1.0 A-Gp [*]	1.0	C-A-C-U-Gp[A]
22	0.9 A-Gp [*]	1.8 A-Up [*] , [*] Gp	1.3 A-Up [*]	0.8 A-Gp [*]	1.0	A-U-A-C-Gp[A]
24	0.4 A-A-Up [*]	1.5 A-A-Up [*] , [*] Gp	1.5 A-A-Up [*] , [*] Gp	0.5 [*] Cp	0.6	C-C-A-A-U-Gp[U]
26	0.9 A-Up [*]	2.9 [*] Gp, [*] Cp (1:2)	1.0 A-Up [*]	2.3 A-Gp [*]	0.9	A-C-C-A-C-C-A-U-Gp[A]
33	0.7 A-Gp [*]	1.3 Up [*]	1.9 A-Up [*]	1.0 A-Gp [*]	1.0	A-U-U-A-Gp[C]
35	2.0 [*] Gp, Up [*]	0.9 [*] Cp	2.0 A-Up [*]	----	1.0	C-A-U-U-Gp[G]
37	1.2 [*] Gp, Cp [*]	----	1.8 A-Up [*]	0.9 [*] Gp, Up [*]	0.6	A-U-U-C-C-Gp[G]
40	1.2 A-A-A-Gp [*]	5.0 [*] Cp, A-A-A-Gp [*] (2:1)	2.0 A-Up [*]	1.9 A-A-A-Gp [*] Up	1.1	C-A-U-U-C-A-A-A-Gp[C]
44	1.8 [*] Gp, Up [*]	----	3.8 Up [*] , Cp [*] (2:1)	----	1.2	C-U-U-U-Gp[G]
54	0.9 A-A-Gp [*]	4.9 A-A-A-A-A-Up [*] A-A-Gp, Up [*] (4:1:2)	6.0 A-A-A-A-A-Up [*] A-A-Gp, Up [*] (2:2:1)	2.3 A-A-Gp [*] , Up [*]	1.0	C-U-C-U-U-A-A-A-A-A-U- U-A-A-Gp[C]
56	0.6 A-Gp [*]	0.6 pppA-Up [*] , [*] Cp	0.6 pppA-Up [*]	2.2 pppA-Up [*] , A-Gp [*]	0.5	pppA-U-C-A-Gp[C]

Table 1. Major leftward RNA of bacteriophage lambda was prepared as described in Methods and elsewhere (6), using three unlabeled and one α -³²P-labeled nucleoside triphosphates at a time. The RNA was digested with RNase T₁ and the resulting oligonucleotides were separated by high voltage paper electrophoresis. Representative fingerprints have been published elsewhere (6) and the numbering system used is shown in Figure 1a. The amount of radioactivity in each oligonucleotide was determined by counting in a low background planchet counter. Molar yields were calculated relative to one mole each of oligonucleotides 19 and 20. Oligonucleotides were eluted from the paper and redigested with pancreatic RNase (0.1 mg/ml, 2 mg/ml yeast carrier RNA) in about 5 μ l of 0.05 M Tris HCl, pH 7.8 for 30 minutes, 37°C. Redigestion products were separated by electrophoresis on DEAE cellulose paper in pyridine acetate at pH 3.5. When the source of the radioactive phosphorus atom could not be deduced unambiguously, the redigestion product was further digested with alkali and the resulting mononucleotides were separated by electrophoresis on Whatman 540 paper at pH 3.5. In all cases, the radioactive phosphate is denoted by an asterisk(*). Brackets [] indicate the 3' nearest neighbor nucleotide. Numbers in parentheses indicate the relative molar yields of labeled redigestion products. These relative yields were determined by counting radioactivity of the appropriate redigestion products and rounding the ratio of counts to the nearest integer. When no numbers are presented the ratio was within $\pm 20\%$ of 1:1. Determination of the exact sequence of oligonucleotide 54 relied upon the presence of the sequences pyAAG[A] and pyAAU[U] in antisense oligonucleotide A46 (identified as the complementary sequence by the presence of 5 Up's) and by the presence of pyAAGAGC[G] (A117) in the pancreatic RNase digest of antisense RNA.

Table 2

PANCREATIC RNase OLIGONUCLEOTIDES DERIVED FROM THE MAJOR LEFTWARD (SENSE) RNA

Oligo-nucleotide number	GTP	ATP	UTP	CTP	Relative Molar Yield	Structure
102	1.1 A-Cp*		1.1 A-Cp*	5.4 A-Cp* (2:1)	1.1 1.1 1.7	A-Cp[G] A-Cp[U] A-Cp[C]
104	1.4 A-A-Cp*	0.8 A*-A-Gp	----	2.0 A-A-Gp*, Cp*	1.0	A-A-G-Cp[C]
105	1.3 Cp*	1.0 Cp*	1.2 Cp*	9.2 Gp*, Cp* (2:1)	1.3 1.0 1.2 2.9	G-Cp[G] G-Cp[A] G-Cp[U] G-Cp[C]
106A	1.3 A-Cp*	2.0 Gp*	----	4.0 A-Cp* (2:1)	1.3 1.3	G-A-Cp[G] G-A-Cp[C]
106B	3.1 A-Cp*	1.9 Cp*	1.0 Cp*	2.6 A-Gp*	1.9 1.0	A-G-Cp[A] A-G-Cp[U]
107	0.8 A-A-A-Cp*	2.4 Gp* A*-A-A-Cp	----	0.7 A-A-A-Gp*	0.8	G-A-A-A-Cp[G]
108	1.2 A-A-A-Gp*	2.0 A*-A-A-Gp, Cp* (2:1)	----	1.1 Gp*	0.9	A-A-A-G-Cp[A]
110	1.0 A-Up*	----	3.0 A-Up* (2:1)	----	1.0 1.0	A-Up[G] A-Up[U]
111	0.7 A-A-Up*	1.0 A*-A-Up	1.0 A-A-Up	----	0.9	A-A-Up[G]
118	1.0 Gp*	----	----	1.8 Gp*, Cp*	0.9	G-G-Cp[C]
119	3.8 A-Gp*, Cp* (2:1)	1.1 Gp*	----	1.4 A-Cp*	1.3	A-G-G-A-Cp[G]
120	----	3.2 A*-A*-A*-A-Up	1.9 A-A-A-A-A-Up	----	0.8	A-A-A-A-A-Up[U]
121	1.0 A-A-Gp*	2.2 Gp*, A*-A-Gp, Cp*	----	1.0 Gp*	0.8	G-A-A-G-Cp[A]
123	2.6 Up*	0.7 Up*	5.0 Gp*, Up* (4:1)	----	2.6 0.7 0.8	G-Up[G] G-Up[A] G-Up[U]
124	----	3.2 Gp*, A-Up* (2:1)	3.0 A-Up* (2:1)	----	1.0 1.0	G-A-Up[A] G-A-Up[U]
126	2.3 A-A-Gp, Up*	1.0 A*-A-Gp	0.7 A-A-Gp*	----	1.0	A-A-G-Up[G]
128	3.4 A-Gp, A-A-Gp* (1:2)	1.9 A-Gp, A*-A-Gp	1.0 Cp*	1.0 Gp*	1.0	A-G-A-A-G-G-Cp[U]
135	1.8 Gp*	0.9 Gp*	2.0 A-Up*	----	0.9	G-G-A-Up[U]
140	3.5 A-A-Gp, Up* (2:1)	2.0 Gp*, A*-A-Gp	1.0 Gp*	----	1.0	G-A-A-G-G-Up[G]
143	3.6 Gp*, Up* (3:1)	----	0.9 Gp*	----	0.9	G-G-G-G-Up[G]
151	3.4 A-A-Gp (2:1), Gp* (3:1)	3.8 A*-A-Gp (2:1), Gp*, Cp* (3:1:1)	----	1.0 Gp*	0.8	G-A-A-G-A-A-G-G-G-Cp[A]
152	----	0.7 pppA-Up	0.8 pppA-Up	0.5 pppA-Up	0.7	pppA-Up[C]

Table 2. RNA preparations analyzed in Table 1 were also digested with pancreatic RNase (Fig. 1b). The resulting purified oligonucleotides were redigested with RNase T₁ (0.1 mg/ml, 2 mg/ml carrier RNA, 30 minutes, 37°C). The presence of a single Gp residue at the 5' end of oligonucleotide 151 was deduced from the structure of partial RNase T₁ digestion products containing only oligonucleotides 19 and 7. The mononucleotides UMP and CMP were not analyzed.

example, one such short RNA contained oligonucleotides 56, 5, 20, 26 and C-A-Gp[G]. The antisense RNA oligonucleotides (see below) were also useful in confirming some of these overlaps. The number of U-Gp oligo-

Figure 2. Sequence of the major leftward RNA of bacteriophage lambda. The oligonucleotides (Tables 1 and 2) were ordered by analysis of partial digestion products of *in vitro* synthesized RNA labeled with ^{32}P -GTP and purified as described in Methods and elsewhere (6). The RNA was digested with low levels of RNase T₁ or carboxymethylated pancreatic RNase (25) and the resulting partial digestion products were fractionated by electrophoresis at pH 3.5 on cellulose acetate, followed by homochromatography (24). Purified partial digestion products were eluted and digested to completion with either RNase T₁ or pancreatic RNase and the resulting oligonucleotides were separated by one dimensional electrophoresis on DEAE cellulose paper in 7% formic acid. Oligonucleotides were identified by their mobilities in this separation plus the products obtained by redigestion with other RNases (pancreatic RNase or RNase T₁, depending on the original digestion). The lines below the sequence indicate some of the partial digestion products obtained by RNase T₁ (straight lines) or carboxymethylated pancreatic RNase (wavy lines).

nucleotides between positions 90 and 96 was deduced from the finding that two G-Up [G] sequences were protected against RNase digestion by hybridization of sense RNA to λdv DNA (carried on $\lambda\text{bio30-7dvl}\text{nin5}$ DNA) (6). Nucleotides beyond position 116 are not present in that DNA (see below).

Antisense RNA.

We have previously shown that p_L and o_L are not transcribed into RNA, and thus must be located somewhere upstream from the natural startpoint, s_L . Sequence studies on antisense RNA were undertaken both to confirm the sequence of the sense RNA and to obtain information about the upstream region.

The template for antisense transcription was a DNA fragment obtained by digestion of λdvl plasmid DNA with the HinII restriction enzyme isolated from Haemophilus influenzae (15) (See Fig. 3). In collaboration with B. Allet, we identified the longest fragment in a HinII digest of EcoRI digested λdvl DNA (about 1500 base pairs long) as the one con-

Figure 3. Genetic and physical map of the immunity region of bacteriophage λ . The 72 to 90 percent segment of the λ papa genome (measured from the left end of the phage) is shown (solid line) with a few important genes and sites positioned approximately to scale. The region between parentheses on the lambda molecule is deleted in λ nin5 strains. A horizontal line above the map is used to indicate the genetic region that was excised from λ vlv2v3 to produce the λ dvl circular plasmid. The novel joint is indicated by a dot on the circle and several sites of cleavage by restriction endonucleases are shown by arrows (see Ref. 8). The Hin fragment containing the novel joint was used as template for antisense RNA indicated by the large arrow. Since one end of the dvl genetic region lies within the $nin5$ deletion, the portion of the circle from parenthesis to dot is missing in $nin5$ strains. Hybridization of antisense RNA with r strands of λ nin5 DNA followed by digestion with RNase was used to purify the sequences on the v2 side of the dot.

taining both the startpoint s_L and, 116 nucleotides away, the novel joint where splicing took place to form the λ dvl plasmid circle. This identification was initially based on the fact that this fragment alone had no identical sized counterpart in a comparable digest of λ wild type DNA. The identification was confirmed by examination of heteroduplexes between the fragment and λ DNA strands in the electron microscope (8). These micrographs showed that the novel joint was located about 150 base pairs from the end of the fragment. Thus only 10% (150/1500) of the DNA comprising the fragment came from the P_L - O_L -major leftward RNA region; the remainder of the DNA came from the area between genes P and Q , which carries no known genes and which is deleted in λ nin5 strains (See Fig. 3).

By DNA-RNA hybridization we found that transcription from the restriction fragment was equally distributed between l and r strands of the fragment. This was rather surprising, since no rightward promoter had previously been observed in the area between genes P and Q in intact DNA (6); apparently, cutting the template introduced at least one new RNA polymerase entry site and startpoint.

The structure of the template fragment made it possible to isolate anti-sense RNA from the p_L-dvl region by hybridization of the RNA to r strands of λ carrying the nin5 deletion (20). This deletion eliminated DNA between genes P and Q, and thus deleted all DNA sequences in the novel joint fragment except those in the p_L-dvl region.

Fingerprints of oligonucleotides obtained by T₁ or pancreatic RNase digestion of antisense RNA labeled with $\alpha^{32}\text{P}$ -GTP are shown in Figures 1c and d. Tables 3 and 4 list the oligonucleotides obtained after digestion of antisense RNA by T₁ RNase and pancreatic RNase, respectively. Oligonucleotides which were expected from the complementary sequence of the major leftward RNA are denoted with crosses (+). Most other oligonucleotides present in antisense RNA were presumably synthesized using the region upstream of the sense RNA startpoint as template. Although we have little information about the relative order of these oligonucleotides in the antisense RNA sequence, our data show that oligonucleotide A14 overlaps the startpoint s_L: this is the only RNase T₁ oligonucleotide in antisense RNA that has (G)-A-U- at its 5' end, which is complementary to the 5' end of sense RNA, pppA-U-C-... Thus from the structure of A14 we conclude that the startpoint is in the center of a sequence [5']-G-A-T-G-T- (on the template strand).

A comparison of oligonucleotides in the antisense RNA which were not predicted from the sense RNA sequence (no crosses in Tables 3 and 4) shows very good agreement with the sequence derived by Maniatis et al (11). These oligonucleotides coming from the region just upstream of the sense RNA startpoint for the operator region are indicated by an underline in the table. One difference between our results and those of Maniatis et al. (11) is -A-C-Gp instead of a -C-C-Gp in oligonucleotide A36. There were also a few oligonucleotides in our catalogue that were not predicted by the sequence of Maniatis et al. (11); these extra sequences may have come from low levels of contamination in our DNA preparations or from partial sequence homologies between the nin region and other parts of DNA. It is interesting to note that the yield of upstream oligonucleotides, based on the published sequence of this region (11), was progressively reduced between s_L and the 5' end of the template.

DISCUSSION

The major leftward operon of bacteriophage lambda codes for a number of important gene products the first of which, gene N, is essential for

Table 3

PANCREATIC RNase OLIGONUCLEOTIDES DERIVED FROM ANTISENSE RNA

Oligo-nucleotide number	GTP	ATP	UTP	CTP	Relative Molar Yield	Structure
A101	3.3 Cp*	11 Cp*	5.0 Cp	n.s.	3.3 11 5.0 ?	<u>Cp[G]</u> ⁺⁺ <u>Cp[A]</u> ⁺⁺⁺⁺⁺ <u>Cp[U]</u> ⁺⁺⁺⁺⁺ Cp[C] ⁺⁺⁺⁺
A102		1.9 A-Cp*	0.5 A-Cp*	8.0 A-Cp*	1.9 0.5 2.8	A-Cp[A] ⁺⁺ A-Cp[U] <u>A-Cp[C]</u> ⁺⁺
A103	0.5 A-A-Cp*	1.0 A*-A-Cp	----	n.s.	0.7	A-A-Cp[G]
A104	0.3 A-A-A-Cp*	1.9 A*-A*-A-Cp	----	0.6 A-A-A-Cp*	0.6	A-A-A-Cp[G]
A105	1.0 Cp*	2.3 Cp*	3.3 Cp*	13.0 Gp*, Cp* (2:1)	1.0 2.3 3.3 6.4	G-Cp[G] ⁺ G-Cp[A] G-Cp[U] ⁺⁺⁺⁺ G-Cp[C] ⁺
A106-A	1.3 A*-Gp, Cp*	----	----	1.4 A-Gp*	0.9	A-G-Cp[G]
A106-B	----	0.7 Gp*	----	1.4 A*-Cp*	0.7	G-A-Cp[C]
A108	0.8 A-A-A*-Gp	2.0 A*-A*-A-Gp	----	1.4 A-A-A-Gp*, Cp*	0.8	A-A-A-G-Cp[C] ⁺
A110	1.8 A-Up*	1.2 A-Up*	6.0 A-Up*	2.4 A-Up*	1.8 1.2 0.3 2.4	A-Up[G] ⁺ A-Up[A] <u>A-Up[U]</u> <u>A-Up[C]</u> ⁺
A111	1.5 A-A-Up*	2.2 A*-A-Up*	2.2 A-A-Up*	----	1.5 0.7	A-A-Up[G] ⁺ A-A-Up[U] ⁺
A112	2.1 Gp*	1.7 Cp*	----	2.6 Gp*	2.0	G-G-Cp[A]
A117	1.5 A-A*-Gp, A*-Gp, Cp*	1.3 A*-A*-Gp*	----	0.8 A-Gp*	0.7	A-A-G-A-G-Cp[G] ⁺
A123	0.8 Up*	2.4 Up*	7.0 Gp*, Up* (6:1)	? Up*	0.8 2.4 1.0 3.0	<u>G-Up[G]</u> G-Up[A] ⁺ G-Up[U] ⁺⁺ G-Up[C] ⁺⁺
A124	3.3 A*-Gp, Up* (2:1)	0.6 Up*	2.0 A-Gp*	----	1.4 0.6	<u>A-G-Up[G]</u> ⁺ <u>A-G-Up[A]</u>
A125	0.8 A-Up*	1.6 Gp*	1.2 A-Up*	----	1.2	G-A-Up[C] ⁺
A126	1.1 A-A-Up*	2.1 A*-A-Up, Gp*	1.0 A-A-Up*	----	1.0	G-A-A-Up[G] ⁺
A129	n.s.	n.s.	0.5 A-A-A-Gp*	n.s.	0.5	A-A-A-G-Up[?]
A131	3.7 A*-Gp, Gp*	----	n.s.	1.3 Gp*	1.3	A-G-G-G-Cp[U] ⁺
A135	5.7 Gp*, Up* (4:1)	0.8 Up*	3.7 Gp*	1.3 Up*	1.7 0.8 1.3	G-G-Up[G] ⁺ <u>G-G-Up[A]</u> G-G-Up[C] ⁺

Table 3. Oligonucleotides produced by T₁ RNase digestion of antisense RNA were separated (Fig. 1c) and analyzed as described in the legend to Table 1. n.s. means that this particular product was not studied. Sequences of oligonucleotides A34, A40, A43, A45, A46, A49 and A50 were deduced, in part, by analysis of the sense RNA sequence (Figure 2), when the nearest neighbor labeling pattern did not permit unambiguous alignment of products. The parentheses in the structure of oligonucleotide A44 indicates that the sequence cannot be unambiguously deduced from our data. A cross (+) indicates oligonucleotides that were predicted from the sequence of the sense RNA. The number of crosses represents the number of times the oligonucleotide was expected to appear. Underlined oligonucleotides are those which are predicted by the sequence (11) for the region immediately upstream from s₁.

Table 4

RNase T₁ OLIGONUCLEOTIDES DERIVED FROM ANTISENSE RNA

Oligo-nucleotide number	GTP	ATP	UTP	CTP	Relative Molar Yield	Structure
A1	2.5 Gp*	2.1 Gp*	4.0 Gp*	2.9 Gp*	2.5	Gp[G] ⁺
					2.1	Gp[A]
					4.0	Gp[U] ⁺⁺
					2.9	Gp[C] ⁺
A2	4.3 Cp*, Gp* (4:1)	0.9 Gp*	1.9 Gp*	0.6 Gp*	0.7	C-G[G]
					0.9	C-Gp[A]
					1.9	C-Gp[U] ⁺⁺
					0.6	C-Gp[C]
A4	0.8 A-Gp*	1.2 A-Gp*, Cp*	----	----	0.7	C-A-Gp[A]
A5	0.4 A-Gp*	0.3 Cp*	n.s.	0.8 Cp*	0.5	C-C-A-Gp[U]
A6	3.2 A-Gp* (2:1)	0.4 A-Gp*	----	1.3 A-Gp*	0.7	A-Gp[G]
					0.4	A-Gp[A]
					1.3	A-Gp[C] ⁺
A7	0.8 A-A-Gp*	1.7 A-A-Gp*	----	----	0.8	A-A-Gp[A]
A8	0.4 A-Gp*	0.4 A-Gp*	----	0.7 A-Gp*, A-Gp*	0.4	A-C-A-Gp[C]
A11	4.6 Up*, Gp* (2:1)	----	----	1.9 Gp*	1.4	U-Gp[G] ⁺
					1.9	U-Gp[C] ⁺
A13	2.5 Up*	0.9 Gp*	2.4 Cp*	1.2 Gp*	0.9	C-U-Gp[A] ⁺
					1.2	C-U-Gp[C] ⁺
A14	1.3 A-Up*	----	2.7 A-Up*, Gp*	----	1.3	A-U-Gp[U] ⁺
A16	1.1 A-Gp*	0.9 Cp*	0.9 A-Gp*	1.1 Up*	0.9	U-C-A-Gp[U] ⁺
A18	0.9 A-A-Up*	1.0 A-A-Up*	0.8 A-A-Up*	1.0 Gp*	1.0	A-A-U-Gp[C] ⁺
A20	0.6 A-Gp*	0.6 Cp*	1.2 A-Gp*, Cp*	0.7 Up*	0.6	C-U-C-A-Gp[U]
A22	0.7 A-A-Up*	0.6 A-A-Up*, Cp*	0.6 A-A-Up*	0.6 Gp*	0.6	C-A-A-U-Gp[C]
A23	0.7 A-A-Up*	1.6 A-A-Up*, Cp*	0.6 A-A-Up*	1.6 Gp*, Cp*	0.8	C-C-A-A-U-Gp[C] ⁺
A30	1.3 Up*	----	0.8 Up*	0.3 Gp*	0.3-1.3	U-U-Gp[C]
A34	1.4 Cp*, Up*	----	2.4 Cp*, Up*, Gp* (2:1:1)	3.4 Up*, Cp*, Gp* (2:1:1)	0.7	U-C-C-U-Gp[C] ⁺ and
					0.7	C-U-U-C-Gp[U] ⁺
A36	0.9 A-A-Cp*	2.3 A-A-Cp*, Up*, Cp*	1.0 A-Up*	1.3 A-A-Cp*, A-Up*, Gp*	0.9	U-A-U-C-A-A-C-Gp[C]
A40	2.2 Cp*, Up*	1.1 Gp*	4.5 Cp*, Up*, Gp* (1:4:1)	1.1 Up*	1.1	U-U-U-C-Gp[U] ⁺
					1.1	C-U-U-U-Gp[A] ⁺
A43	1.1 Up*	----	2.7 Cp*, Up* (2:1)	2.7 Up*, Cp*, Gp*	1.0	C-C-U-U-C-U-Gp[C] ⁺
A44	0.5 A-Up*	2.4 Up*	2.3 A-Up*, Up*, Gp*	----	0.5	(U-A-U-) (U-)U-A-U-G[U]
A45	0.7 A-A-A-Gp*	3.4 A-Gp*, Up*, Cp*	n.s.	4.9 A-Up*, A-Cp*	0.7	U-A-U-C-A-C-A-C-A-C-
				A-Cp*, Cp*, A-A-A-Gp*		C-C-C-A-A-A-Gp[C] ⁺
A46	1.2 A-A-Gp*	4.5 Up*, A-A-Up* (2:1:2)	6.8 Up*, Cp*, A-A-Up* (4:1:1)	----	1.2	C-U-U-A-A-U-U-U-U-U-
						A-A-Gp[A] ⁺
A49	2.1 A-Up*, Gp*	2.9 Cp*	4.5 Up*, Cp*	4.0 Up*, A-Cp*	1.0	U-C-A-C-C-U-U-C-A-U-
						Gp[G] ⁺
A50	2.2 A-Gp*	1.9 Cp*	3.3 Up*, Cp*	4.7 Up*, Cp*	1.0	C-C-C-U-U-U-C-U-U-C-A-
						Gp[G] ⁺

Table 4. Antisense RNA was digested with pancreatic RNase (Fig. 4d) and analyzed as described in the legends to Tables 1, 2 and 3.

normal phage growth. An interesting characteristic of the operon is that neither promoter nor operator is transcribed into RNA (6). Instead, transcription starts both *in vivo* and *in vitro* at a startpoint, s_L, located downstream from these elements. We have determined the sequence of 149 nucleotides downstream of s_L and obtained some sequence information upstream of s_L.

The Beginning of the Gene N Cistron.

The initial sequence of the major leftward RNA beginning at startpoint ϕ_1 is shown in Figure 2 and, in a clearer form, in Figure 4. There are eight AUG or GUG codons which should be considered as possible initiators of the N protein. These are located at positions 24, 30, 90, 92, 94, 122, 145 and 147. At present we are unable to determine which of these, if any, serves as the initiator of protein synthesis in vivo. We note, however, that four of them, at positions 24, 30, 92 and 94, are unlikely candidates since they are followed by terminators in phase.

The first initiation codon not followed by an in-phase terminator is located at position 90. An extensive untranslated leader region is not unusual in messenger RNAs. The first mRNA to be sequenced from its 5' end to the initiation codon, Q β RNA, had a segment of 61 nucleotides before the protein initiator (21). The mRNA of the trp operon of E. coli also has a long leader sequence (26), whereas the gal (27) and lac (28) operons contain the initiation AUG codons within thirty nucleotides of their 5' ends.

Figure 4. Sites in the major leftward RNA of bacteriophage lambda. The nucleotides shown above the sequence indicate changes that were observed in the sequence with λ imm434b2 DNA as template. The exact nature of the change at positions 22 or 23 is unknown. λ dv1 indicates the novel joint in the λ dv1 plasmid. ϕ_1 is the 5' end of the relatively stable 4.5S λ RNA (29). This RNA has pG- at its 5' end and extends from position 89 to about 190. "HpaII" and "HaeIII" indicate positions we propose to be cleavage sites for these restriction endonucleases.

Initial Sequence of " ϕ_1 "

Several years ago, Marcaud et al. (29) reported that a 4.5S λ specific RNA which they called " ϕ_1 " accumulates after induction of a λ lysogen. We

have reported that this RNA is a segment of p_L RNA beginning near position 90 (6). In collaboration with H. Lozeron and W. Szybalski we have determined that the 5' end of l_1 is the pG- at position 89 (Figure 4) and that the RNA extends to about position 190 (18, 19, 30). This RNA fragment is a relatively long-lived intermediate in the in vivo degradation of p_L RNA (18, 19, 30).

Position of the point of Cyclization of λ dv1

The circular DNA molecule of λ dv1 comprises the segment from 73.25 to 84.3% on the λ DNA map measured from the left end (8, 20). The point of cyclization or "novel joint" is located slightly to the left of s_L on the λ map and thus serves as an important reference point for measuring this area of the chromosome (See Fig. 3).

The novel joint has been determined to be between positions 116 and 117 of the sense RNA transcript (Figure 4), i.e., beginning at s_L and proceeding leftward the first base pair which differs between the DNAs of λ dv1 and λ is at position 117. It represents a change from a C in λ to a purine in λ dv1. This conclusion follows from analysis of antisense RNA templated from the λ dv1 DNA fragment. In this RNA all oligonucleotides predicted by the sense sequence up to position 116 are present and those beyond 116 are absent. For example, between positions 120 and 109 the complementary RNA sequence would be -U-A-C-G-G-C-C-A-A-U-G-C- predicting the presence of RNase T_1 oligonucleotide [G]C-C-A-A-U-Gp[C] and the pancreatic RNase oligonucleotide [Py]A-Cp[G] in antisense RNA. The catalogue of antisense RNA templated from λ dv1 DNA contained the former oligonucleotide (A23) but not the latter, and thus the novel sequence in this region must start between positions 117 and 120. Moreover the only pancreatic oligonucleotide in the antisense RNA which could overlap the 5' end of A23 is A105, [Py]G-Cp[C], and the predicted overlapping oligonucleotide, complementary to λ sense RNA, [Py]G-G-Cp[C], is missing. Thus the sequence of λ and λ dv1 DNA diverge at position 117. This location of the dv1 novel joint confirms our earlier conclusion about its position which was based upon protection of sense RNA against RNase digestion when the RNA was hybridized to λ dv1 DNA (6).

Restriction Endonuclease Sites.

The RNA sequence shown in Figure 4 contains two sequences which probably correspond to restriction nuclease sites in the DNA template. Murray and Morrison observed that the HaeIII endonuclease hydrolyzes DNA at (5')-G-G-C-C-

(3') sequences (31). We find such a sequence at positions 114-117. Allet observed that λ DNA, but not λ dv1 DNA, is cut in this neighborhood by the HaeIII restriction enzyme (32 and B. Allet, pers. comm.). This observation is in agreement with the location for the λ dv1 endpoint discussed above since the last nucleotide of the site is changed in λ dv1.

Boyer and his collaborators have proposed that some instances of illegitimate recombination may result from cutting and splicing of two DNA molecules at identical restriction enzyme target sites (33). Interestingly, the λ dv1 endpoint falls within such a site (G-G-C-C). However, formation of the novel joint eliminated the site. Thus the illegitimate recombination event leading to formation of λ dv1 must have been more complex than the simple cutting and rejoining of DNA at two-G-G-C-C- sequences. Moreover, there is no indication that E. coli contains a restriction enzyme recognizing this sequence.

A second palindromic sequence, -U-C-C-G-G-A- is at positions 129-134. We propose (as cited in reference 10) that this is the HpaII cleavage site which Allet identified as being about 150-160 nucleotides from the HinIII cleavage site in the σ_L operator (34), i.e., about 120-130 nucleotides from S_L .

Sequence Changes Using Different Templates

In a previous communication we compared fingerprints of p_L RNA made from DNA templates carrying cis-acting mutations affecting left operon transcription. We concluded that promoter and operator were not transcribed since mutants sex1 and v2 did not alter the RNA sequence (6). In addition, only four scattered single base changes were observed with the promoter-operator substitution hybrid λ imm434. These changes are specified in Figure 4. We presume that they were derived from phage 434 when the λ imm434 hybrid was constructed. They have no known phenotypic effect.

The one difference in the untranscribed region between our data and that of Maniatis et al. (11) is in oligonucleotide A36, in which we find U-A-U-C-A-A-C-Gp instead of U-A-U-C-A-C-C-Gp. We presume the reason for this difference is that our template DNA, derived from λ dv1, contains the v2 mutation in the operator although it is possible that this change could merely reflect nonisogenic backgrounds of the DNAs. It is interesting to note that the change decreases two forms of operator symmetry while, as pointed out by W. Szybalski, it increases a third, relative to the wild-type sequence. The v101 operator mutation has the same effects on the symmetries

of this region, although the point of mutation is five nucleotides removed from the v2 change (11). These changes are illustrated in Figure 5. Hu and Szybalski have indications that v2 is simultaneously an operator-constitutive and a promoter-up mutation (personal communication); they postulate that such a dual effect could be accounted for by the fact that more than one form of symmetry is affected.

Figure 5. Relationship between p_L RNA sequence, antisense RNA catalogue, and repressor binding site sequence. The sequence immediately downstream of the Hin II restriction site, as determined by Maniatis *et al.* is shown as a DNA duplex. The p_L RNA and anti- p_L RNA transcripts are aligned based on the correspondences observed. This defines the positions of s_L in the DNA sequence, and permits the conclusion that the distance from the Hin II site to s_L is 34 nucleotide pairs. In this 34 nucleotide upstream area our antisense oligonucleotide catalogue is arranged to correspond with the DNA sequence. The sequence in this region is slightly changed in two strains which carry the operator oL mutations, $v101$ and $v2$ as determined by Maniatis *et al.* and ourselves, respectively. The boxes at the bottom of the figure indicate three base pairing symmetries in the wild type sequence with stars at their centers. Small arrows (up or down) indicate whether the base changes resulting from the $v2$ and $v101$ mutations strengthen or weaken each symmetry. Both mutations decrease two forms of symmetry and increase a third symmetry in this region. Bars within the DNA sequence are used to indicate two segments containing exclusively A·T or G·C base pairs. Each of the operator mutations also reduces the compositional bias of one of these regions.

Another possibility is that such mutations act by altering the ease with which RNA polymerase could traverse the interval between promoter and startsite. As we pointed out previously (35) blocks of nucleotide with extreme biases in base compositions could produce such a hindrance. It is interesting to note in Figure 5 that each operator mutation reduces the compositional bias of such a block.

ACKNOWLEDGEMENTS

We thank Barbara P. Stark and Pat A. Twose for excellent technical help and W. Szybalski for drafting Figure 5 and for his enthusiasm, advice, critical reading of the manuscript, and material support of F.R.B. This work was supported by grants GB-32152X from NSF (to J.E.D.) and CA-07175 from NCI and GB 2096 from NSF (to W. Szybalski). J.E.D. holds Research Career Development Award 5-K04-32770 from NIH.

#McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, WI 53706. Present address: Department of Genetics, University of Wisconsin, Madison, WI 53706, USA

REFERENCES

- 1 Jacob, F., and Monod, J. (1961) J. Mol. Biol. 3, 318-356
- 2 Epstein, W., and Beckwith, J.R. (1968) Ann. Rev. Biochem. 37, 411-436
- 3 Chamberlin, M. Ann. Rev. Biochem. 43, 721-775
- 4 Ptashne, M. (1971) in The Bacteriophage Lambda (edit. by Hershey, A.D.), 221-237. (Cold Spring Harbor Laboratory, New York)
- 5 Szybalski, W., Bovre, K., Fiantdt, M., Guha, A., Hradecna, Z., Kumar, S., Lozeron, H.A., Maher, V.M., Nijkamp, H.J.J., Summers, W.C., and Taylor, K. (1969) J. Cell. Physiol., 74 (Suppl. 1), 33-70
- 6 Blattner, F.R., and Dahlberg, J. . (1972) Nature New Biology 237, 227-232.
- 7 Maniatis, T., and Ptashne, M. (1973) Nature 246, 133-136
- 8 Blattner, F.R., Fiantdt, M., Hass, K.K., Twose, P.A., and Szybalski, W. (1974) Virology 62, 458-471
- 9 Maurer, R., Maniatis, T., and Ptashne, M. (1974) Nature 249, 221-223
- 10 Maniatis, T., Ptashne, M., and Maurer, R. (1973) Cold Spring Harbor Symp. Quant. Biol. 38, 857-867
- 11 Maniatis, T., Ptashne, M., Barrell, B.G., and Donelson, J. (1974) Nature 250, 394-397
- 12 Blattner, F.R., Dahlberg, J.E., Boettiger, J.K., Fiantdt, M., and Szybalski, W. (1972) Nature New Biol. 237, 232-236
- 13 Vapnek, D., and Rupp, W.D. (1971) J. Mol. Biol. 60, 413-424
- 14 Helling, R.B., Goodman, H.M. and Boyer, H.W. (1974) J. Virol. 14, 1235-1244
- 15 Kelly, jun., T.J., and Smith, H.O. (1970) J. Mol. Biol. 51, 393-409
- 16 Chamberlin, M. and Berg, P. (1962) Proc. U.S. Nat. Acad. Sci. 48, 81-94
- 17 Symons, R.H. (1969) Biochim. Biophys. Acta 190, 548-550
- 18 Lozeron, H.A., Funderburgh, M.L., Dahlberg, J.E., Stark, B.P., and Szybalski, W. (1972) Abstracts Ann. Meet. Amer. Soc. Microbiol. 237,
- 19 Lozeron, H.A., Szybalski, W., and Dahlberg, J.E. Manuscript in preparation
- 20 Fiantdt, M. Hradecna, Z., Lozeron, H.A., and Szybalski, W. (1971) in The Bacteriophage Lambda (edit. by Hershey, A.D.), 329-354, Cold Spring Harbor Laboratory, New York
- 21 Billeter, M.A., Dahlberg, J.E., Goodman, H.M., Hindley, J., and Weissmann, C. (1969) Cold Spring Harbor Symp. Quant. Biol. 34, 635-645
- 22 Lebowitz, P., Weissman, S.M., and Radding, C.M. (1971) J. Biol. Chem. 246, 5120-5139
- 23 Sanger, F., Brownlee, G.G., and Barrell, B.G. (1965) J. Mol. Biol. 13, 373-398

- 24 Brownlee, G.G. and Sanger, F. (1969) Eur. J. Biochem. 11, 395-399
- 25 Contreras, R. and Fiers, R.W. (1971) FEBS Lett. 16, 281-290
- 26 Bronson, M.J., Squires, C., and Yanofsky, C. (1973) Proc. U.S. Nat. Acad. Sci. 70, 2335-2340
- 27 Musso, R., Sklar, J., Yot, P., de Crombrughe, B., Pastan, I. and Weissman, S. (In Press) Proc. U.S. Nat. Acad. Sci.
- 28 Maizels, N.M. (1973) Proc. U.S. Nat. Acad. Sci. 70, 3585-3589
- 29 Marcaud, L., Portier, M.M., Kourilsky, P., Barrell, B.G. and Gros, F. (1971) J. Mol. Biol. 57, 247-261
- 30 Lozeron, H.A., Dahlberg, J.E., and Szybalski, W. Manuscript in preparation
- 31 Murray, K. and Morrison, A. In preparation
- 32 Allet, B. and Solem, R. (1974) J. Mol. Biol. 85, 475-484
- 33 Roulland-Dussoix, D., Yoshimori, R., Greene, P., Betlach, M., Goodnan, H.M., and Boyer, H.W. (1974) In "Microbiology 1974", David Schlessinger, Ed. A.S.M. Press, Washington
- 34 Allet, B. (1973) Biochemistry 12, 3972-3977
- 35 Blattner F. In "Molecular Cytogenetics", B.A. Hamkalo and J. Papaconstantinou, Eds. Plenum Publishing Corporation, New York, 293-300