Proc. Nat. Acad. Sci. USA
Vol. 71, No. 12, pp. 4940–4944, December 1974

The 5′-Terminal Nucleotide Sequence of Galactose Messenger Ribonucleic Acid of *Escherichia coli*

(galactose operon/UDPgalactose 4-epimerase/regulation/transcription/RNA sequences)

RICHARD E. MUSSO*, BENOIT DE CROMBRUGGHE*, IRA PASTAN*, JEFFREY SKLAR†, PIERRE YOT†‡, AND SHERMAN WEISSMAN†

* Laboratory of Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20014; and
† Departments of Human Genetics, Molecular Biophysics and Biochemistry, and Medicine, Yale University School of Medicine, New Haven, Connecticut 06511

Communicated by DeWitt Stetten, Jr., August 7, 1974

ABSTRACT **The 5′-terminal sequence of mRNA from the galactose operon of *E. coli* has been determined. *gal* RNA is synthesized *in vitro* by means of a kinetically controlled purified transcription system and is isolated by a two-step RNA·DNA hybridization scheme. The following sequence of the first 77 nucleotides has been deduced by analysis of oligonucleotides produced by digestion with T_1, pancreatic, and carboxymethylated pancreatic ribonucleases: pppA-U-A-C-C-A-U-A-A-G-C-C-U-A-A-U-G-G-A-G-C-G-A-A-U-U-A-U-G-A-G-A-G-U-U-C-U-G-G-U-U-A-C-C-G-G-U-G-G-U-A-G-C-G-G-U-U-A-C-A-U-U-G-G-A-A-G-U-C-A-U-A-C-C-U-G-U. Residues 27–77 correspond to the amino terminal 17 amino acids of UDPgalactose 4-epimerase (EC 5.1.3.2), the protein specified by the promoter-proximal structural gene of the operon. A self-complementary sequence occurs near the 5′ terminus; 12 of 15 nucleotides between residues 4 and 18 are symmetrically located about position 11. The sequence of residues 22–33 closely resembles part of the lactose operator sequence reported previously [Gilbert, W. & Maxam, A. (1973) *Proc. Nat. Acad. Sci. USA* 70, 3581–3584].**

The *gal* operon of *Escherichia coli* contains the structural genes *E*, *T*, and *K* which specify the enzymes UDPgalactose 4-epimerase (EC 5.1.3.2, referred to as epimerase) galactose 1-phosphate uridylyltransferase (EC 2.7.7.10), and galactokinase (EC 2.7.1.6), respectively (1, 2). The *E* cistron is preceded by a regulatory region containing promoter and operator sites (2, 3). Expression of this operon is subject to two modes of regulation; each involves control at the transcriptional level and has been demonstrated *in vitro*. The *gal* repressor protein interacts with the *gal* operator locus to exert a negative control (2, 4–9). Positive control is provided by cyclic AMP and its receptor protein, CRP, which enable RNA polymerase to form a preinitiation complex at the *gal* promoter (10–12). An additional system of negative control has been proposed (13) but not demonstrated *in vitro*. There appears to be a functional overlapping between the regulatory sites for the *gal* operon (9).

Knowledge of the nucleotide sequences of these DNA sites is a prerequisite for a detailed understanding of *gal* regulation. A comparison with similar sequences of other bacterial or phage operons could aid in elucidating general features of DNA regulatory loci. For these reasons we have undertaken a study of the nucleotide sequence of the *gal* regulatory region.

Our initial approach has been to examine the 5′-terminal sequence of *gal* mRNA. This sequence should define the initiation sites for transcription and translation. It should also contain the ribosome binding site for the *E* cistron and possibly part of the operator–promoter region. Our results, presented below, provide a 5′-terminal sequence for a bacterial mRNA initiated from a wild-type promoter.

MATERIALS AND METHODS

Materials. α-^{32}P-Labeled nucleoside triphosphates (NTPs) (specific activity >80 Ci/mmol) were purchased from New England Nuclear Corp.; other reagents were obtained from various commercial sources. The bacteriophage strains

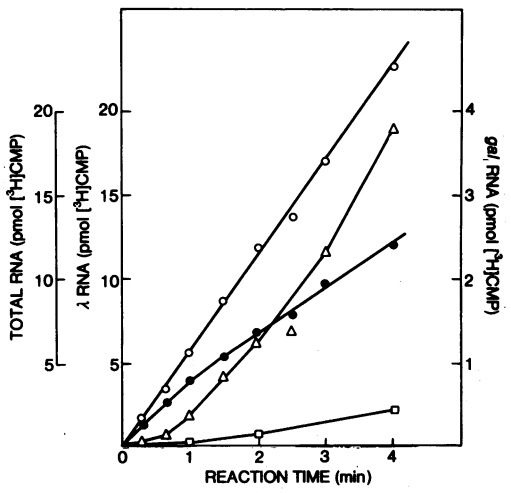

FIG. 1. Kinetics of *gal* RNA synthesis. Transcription of λp*gal*8 DNA was allowed to proceed at 20° for the indicated time as follows. RNA polymerase (25 μg/ml) was preincubated with λp*gal*8 DNA (47 μg/ml) in the presence of 20 mM Tris·HCl at pH 7.9, 100 mM KCl, 0.1 mM dithiothreitol, 15 μg/ml of CRP, 0.1 mM cyclic AMP, 40 μM each of ATP, GTP, UTP, and [δ-^{3}H]-CTP (15 Ci/mmol), and 4 mM EDTA. After 4 min at 20° RNA synthesis was initiated by addition of $MgCl_2$ to 15 mM and 10 μg/ml of rifampicin. The RNA was collected by ethanol precipitation and analyzed as described in *Materials and Methods*. Total RNA precipitable by 5% trichloroacetic acid (○), RNA hybridizable to λ DNA filters (●), and *gal*-specific RNA (△) made in the presence of cyclic AMP are indicated. Apparent *gal* RNA made in the absence of cyclic AMP (□) is negligible.

Abbreviations: CRP, cyclic AMP receptor protein; P-RNase, bovine pancreatic RNase A; CM-RNase, ε-lysine 41 carboxymethylated pancreatic RNase A; NTP, nucleoside 5′-triphosphate.

‡ Present address: Institut de Biologie Moléculaire, Université de Paris, Paris, France.

TABLE 1. *Analysis of oligonucleotides obtained by complete digestion with T_1 RNase*

No.	P-RNase digestion products from the [α-^{32}P] NTP precursor: G	A	C	U	Sequence deduced
T23		G		G	G[A,U]
T22	AG[a]	AG	AG	AG	AG[A,C,U]
T21	C,G	G			CG[G,A]
T27	U,G				UG[G]
T20	AAG[a]	AAG[a]		AAG	AAG[U]
T26	AG[a]	U	AG		UAG[C]
T16	U,G		U	U,C	U(U,C)UG[G][b]
T14	C,G	U	AC[C][a]	U	UUACCG[G]
T10	AAU	AAU[a]	C	AAU[a]	CCUAAUG[G]
	G	U		C	
T8	AU	AAU[a]		AAU[U][a]	AAUUAUG[A]
		U,G		AU	
T7	U,G	U,AC	AC[a]	AU[U][a]	U(U,ACAU)UG[G][c]
				U	
T5	U	C,AU	AC[C][a]	AU[a]	UCAUACCUG[U]
			U	C,G	
T33	AAG	AAG[a]	AAG	pppAU[a,d]	pppAU(ACC,AU)AAG[C]
		C,AU	AC[C][a]	AU[a]	
		pppAU[A][d]			

The T_1 oligonucleotides were digested with P-RNase and products fractionated on DEAE paper by electrophoresis at pH 3.5. Products were identified by their mobilities. Nearest neighbors are indicated in brackets where necessary. Underscoring indicates nearest integral molar yield in excess of one.

[a] Base analysis was performed and nearest neighbor is indicated where determined.

[b] The sequence of T16 was determined as U-U-C-U-G-[G] by the product U-U-C-[U] produced by P-RNase cleavage after modification of [α-^{32}P]UTP-labeled T16 by the carbodiimide CMCT.

[c] The sequence of T7 was determined as U-U-A-C-A-U-U-G-[G] by the products U-U-A, C-A, and U-U-G obtained after U_2 RNase digestion of T7 labeled by [α-^{32}P]UTP.

[d] The product pppA-U was identified by its low mobility on DEAE paper at both pH 3.5 and pH 1.7.

TABLE 2. *Analysis of oligonucleotides obtained by complete digestion with P-RNase*

No.	T_1 RNase digestion products from the [α-^{32}P] NTP precursor: G	A	C	U	Sequence deduced
P1a	[a]	[a]		C[a]	C[A,U,G]
P1b		[a]	AC[C][b]		AC[A,C]
P5a	C,AG[b]		AG		AGC[G]
P5	AAG[b]	AAG[b]	AAG		AAGC[C]
			C		
P15	C,AG[b]	G	AG		GGAGC[G]
	G				
P7	U	U	U		U[A,G,C]
P4	AU	AU		AU[U][b]	AU[A,G,U]
P11		pppAU[A][b]		pppAU[b]	pppAU[A]
P2	AAU	AAU[b]		AAU[b]	AAU[G]
P14	G,U	U		G,U	GGU[U,G,A]
P6		G,AAU[b]		AAU[U][b]	GAAU[U]
P10a[c]	AG[b]	G,AG		AG,U	GAGAGU[U]
P10b[c]	AAG[b]	AAG[b]	U	AAG	GGAAGU[C]
	G	G			

The products obtained after T_1 RNase digestion of P-oligonucleotides were analyzed as described in the legend to Table 1.

[a] C and A-C were not transferred from the Cellogel strip to the DEAE-cellulose thin-layer plate in mapping the primary digest by P-RNase. Their presence was indicated by analysis of T_1 oligonucleotides and confirmed for [α-^{32}P]CTP- and [α-^{32}P]UTP-labeled RNA digests.

[b] Analyzed by base hydrolysis.

[c] P10a and P10b comigrated on the primary map, but P10a was clearly the predominant component in *gal* RNA synthesized for 40 sec. The individual sequences of P10a and P10b were confirmed when the oligonucleotides were obtained from separate CM-RNase partial digestion fragments.

λ*cI857S7* and λp*gal8cI857S7* were prepared by infection and banded in cesium chloride density gradients. DNA was extracted and separated strands were prepared as previously described (12). CRP was purified as described by Anderson *et al.* (14). *E. coli* RNA polymerase holoenzyme was prepared through step 5 of the method of Berg *et al.* (15). Carboxymethylated pancreatic RNase (CM-RNase) prepared by the method of Heinrikson (16) was the gift of Dr. P. Lebowitz.

Preparation of gal RNA. In vitro transcription of λp*gal8*-DNA was controlled kinetically as described by Blattner and Dahlberg (17) with slight modifications as indicated in the *text*. After the desired time the reaction was terminated by adding EDTA to 50 mM, sodium dodecyl sulfate to 0.2%, 100 μg of tRNA, and 1 volume of neutralized phenol. The phenol-extracted RNA was separated from mononucleotides on a 0.9 × 60-cm column of Sephadex G-50 Fine (eluted with 10 mM Tris·HCl at pH 7.9, 10 mM $MgCl_2$, 1 mM EDTA), treated with 15 μg/ml of DNase for 10 min at 0°, and again extracted with phenol in the presence of dodecyl sulfate. After precipitation with 2 volumes of ethanol for 10 hr at −20°, the RNA was collected by centrifugation (20 min at 10,000 × *g*), dried *in vacuo*, and resuspended in 2 × SSC (0.03 M sodium citrate, 0.3 M sodium chloride, pH 7). λ DNA was immobilized on nitrocellulose filters as described by Gillespie and Spiegelman (18) and hybridizations were performed in 2 × SSC at 65°. The total RNA was first hybridized against at least a 100-fold excess of λ DNA on filters for 24 hr to remove λ RNA. The unhybridized RNA (about 20% of the total) was hybridized to a 50-fold molar excess of λp*gal8 l* strand in liquid phase for 5 hr. The *gal* RNA·DNA hybrids (about 6% of total RNA) were collected on nitrocellulose filters and then treated with T_1 RNase (5 U/ml in 50 mM Tris·HCl at pH 7.4, 2 mM EDTA) or bovine pancreatic RNase A (P-RNase) (0.2 μg/ml in 2 × SSC) for 30 min at 23°, and the residual RNase was inactivated with sodium iodoacetate for 45 min at 54° (19). The *gal* RNA was eluted in water for 4 min at 80–90°, treated for 5 min at 37° with 50 μg/ml of DNase in 20 mM Tris·HCl at pH 7.4, 10 mM $MgCl_2$, extracted with phenol, and precipitated by ethanol. This RNA (60%–70%) hybridized to λp*gal8 l* strand and less than 5% to λ *l*, λ *r*, or λp*gal8 r* strands; it was suitable for sequence analysis.

RNA Sequence Analyses. Sequence analyses were performed as described elsewhere (20, 21). Primary digests by T_1 RNase and by P-RNase were prepared for *gal* RNA separately labeled by each of the four α-^{32}P-labeled nucleoside triphosphates, and the oligonucleotides were fractionated on two-dimensional maps (pH 3.5 electrophoresis on Cellogel followed by homochromatography on DEAE-cellulose). Larger fragments were obtained by partial digestion with CM-RNase (23).

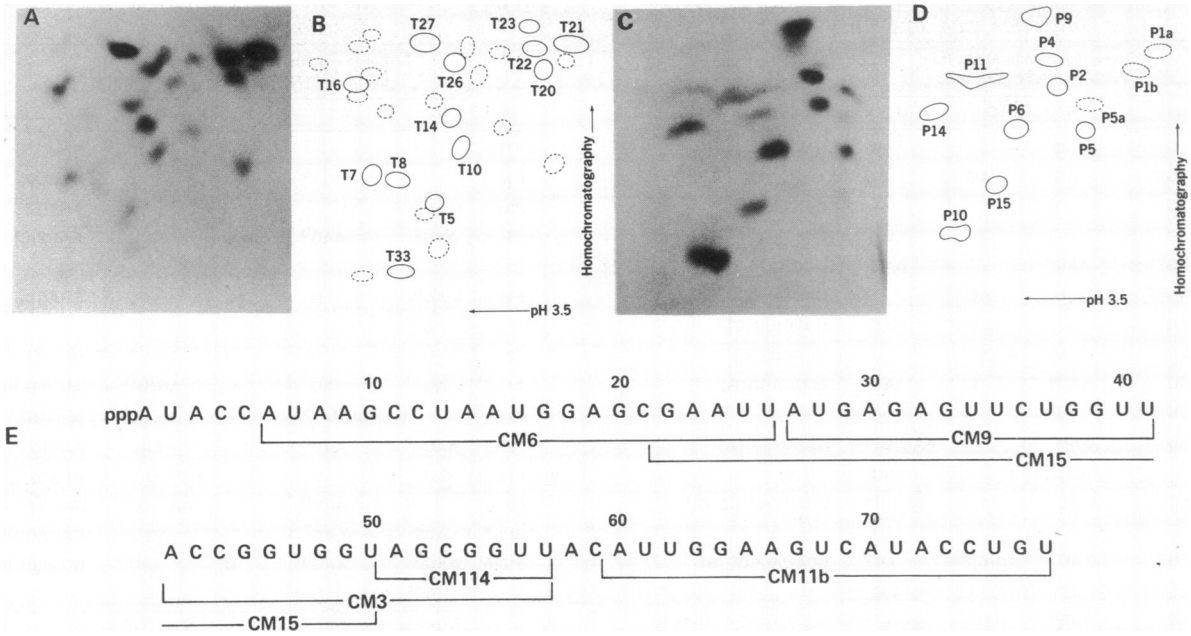

FIG. 2. RNase digestion products from *gal* RNA. (A) *gal* RNA labeled by [α-^{32}P]GTP in a 40-sec reaction at high nucleoside triphosphate concentrations (150 μM) was digested with 25 units of T_1 RNase in 5 μl containing 200 μg of carrier tRNA. After electrophoresis (right to left), the region between the xylene cyanol FF and methyl orange markers was transferred to a DEAE-cellulose thin-layer plate for homochromatography (bottom to top). (B) Tracing of (A) where solid circles indicate oligonucleotides obtained from a 40-sec reaction at low substrate concentration (40 μM) and broken circles indicate T_1 oligonucleotides which appear in transcripts from 60- to 100-sec reactions at low substrate concentrations. (C) *gal* RNA labeled by [α-^{32}P]PTP in a 40-sec reaction at low substrate concentrations (40 μM) was digested with 7 μg of P-RNase in 7 μl containing 200 μg tRNA. (D) Tracing of (C) with solid circles indicating major products labeled by [α-^{32}P]ATP and broken circles indicating other major P-RNase oligonucleotides detected with different α-^{32}P-labeled precursors (5a) or not transferred in the routine procedures (1a and 1b). (E) Bars indicate the CM-RNase partial digestion fragments important for determining the 5′ terminal sequence of *gal* mRNA. Symbols for the phosphodiester bonds linking all residues have been omitted for clarity.

RESULTS

Kinetics of gal RNA Synthesis. *gal* RNA synthesis depends on the presence of cyclic AMP in our *in vitro* system (Fig. 1). Although there is no lag before onset of synthesis of either total RNA or λ-specific RNA, the full rate of *gal* RNA production is attained only after 30 sec. For RNA sequence studies, reactions were incubated 30–100 sec with a progressive increase in the size of the *gal* transcript. The sequence analyses discussed below correspond to a 40-sec reaction under the conditions indicated.

T_1 RNase Digest. A two-dimensional map of *gal* RNA after T_1 RNase digestion is shown in Fig. 2A. Secondary digestion with P-RNase (Table 1) provided the sequence for each T_1 oligonucleotide except T33, T16, and T7. The sequence of the latter was determined by U_2 RNase cleavage; T16 was analyzed by P-RNase after *N*-cyclohexyl-*N*′-[β-morpholinyl-(4)-ethyl] carbodiimide methyl-*p*-toluene sulfonate (CMCT) modification (22). The sequences of these T_1 oligonucleotides were confirmed and the ambiguity of T33 resolved by analysis of CM-RNase partial digestion fragments (see below).

P-RNase Digest. The map of a typical P-RNase digest of *gal* RNA is shown in Fig. 2C. Analyses with T_1 RNase (Table 2) established all sequences. The two oligonucleotides (10a and 10b) that co-migrated in the P-RNase fingerprint were resolved in separate CM-RNase fragments and their individual sequences were confirmed (see below).

CM-RNase Partial Digest. The T_1 and pancreatic oligonucleotides were ordered by analysis of fragments from partial digests with CM-RNase (Table 3 and Fig. 2E). The sequences of T16 and T7 were confirmed by analyses of CM9 and CM114. The ambiguity in T33 was resolved by the sequence of CM6. The fragments P10a and P10b were resolved in CM9 and CM11b, respectively. Fragments CM8 and CM15 in addition to those already mentioned allowed the complete sequence of the first 77 nucleotides to be assembled.

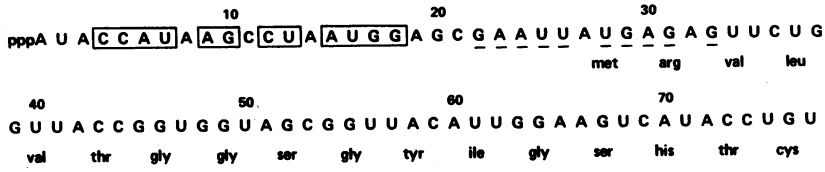

FIG. 3. Correlation of *gal* mRNA and UDPgalactose 4-epimerase sequences. A self-complementary sequence in the RNA is indicated by the boxed residues. The sequence of residues 22–33 is nearly identical (as underlined) to part of the lactose operator sequence.

TABLE 3. *Analysis of major oligonucleotides obtained by limited digestion with CM-RNase*

Oligonucleotide number	[α-32P]NTP Precursor	Products of Extensive T_1 RNase Digestion of Partial Product	Relative Mobility of T_1 RNase Product	Products of Extensive P-RNase Digestion of Partial Products	Relative mobility of P-RNase Product[a]	Sequence Deduced
CM6	ATP	CCUAAUG	0.135			AUAAGCCUAAUGG
		AUAAG	0.32			AGCGAAUU(A)
		AAUU(A)	0.35			
		CG(A)	1.65			
		G(A)	1.78			
	CTP+UTP	CCUAAUG	0.133	GGAGC	0.147	
		AUAAG(C)	0.33	GAAU	0.39	
		AAUU	0.39	AAU	1.02	
		AG(C)	1.46	AAGC(C)	1.05	
				AU[b]	1.44	
	UTP	CCUAAUG	0.132			
		AUAAG	0.32			
		AAUU	0.36			
CM8	CTP+UTP	UAG(C)	0.63	GGU(U)	0.224	ACCGGUGGUAGCG
		UU	0.78	AGC	1.18	UU
		G(U)	1.80	AC	1.75	
		ACCG	1.20			
	UTP	UU	0.81	GGU(U)	0.272	
		G(U)	1.77			
CM9	GTP	UUCUG	0.102			AUGAGAGUUCUGG
		AUG	0.61			UU(A)
		AG	1.47			
	ATP	AUG	0.59			
		UU(A)	0.82			
		AG(A)	1.44			
	CTP	UUCUG	0.098			
	UTP	UUCUG	0.106	GAGAGU(U)	0.054	
		AUG	0.57	GGU(U)	0.27	
		UU	0.82	AU	1.62	
		AG(U)	1.45			
		G(U)	1.77			
CM11b	CTP+UTP	UCAUACCUG	0.055	GGAAGU	0.041	AUUGGAAGUCAUAC
		AUUG	0.193	GU	0.74	CUGU
		AAG	1.20	AU	1.41	
				AC	1.77	
CM15	CTP+UTP	AAUUAUG	0.047	GAGAGU	0.038	GAAUUAUGAGAGUU
		UUCUG	0.108	GGU	0.23	CUGGUUACCGGUG
		UUACCG	0.189	GAAU	0.39	GU
		AG	1.23	AU	1.57	
		G	1.51	AC	2.06	
CM114	GTP	AG	1.49			AGCGGUU(A)
		CG(G)	1.65			
	ATP	UU(A)	0.81			
	CTP	AG(C)	1.46	AGC	1.39	
	UTP	UU	0.80	GGU(U)	0.29	
		G(U)	1.85			

Oligonucleotides obtained from partial digestion of *gal* mRNA by CM-RNase were digested to completion with T_1 RNase or P-RNase and the products were fractionated on DEAE paper by electrophoresis at pH 1.7. The identity of these products was confirmed by treating them with the reciprocal enzyme and analyzing the mobilities of the resulting products by electrophoresis at pH 3.5 on DEAE paper. Underscoring indicates nearest integral molar yield in excess of one.

[a] Relative mobilities of T_1 RNase and P-RNase products on DEAE paper at pH 1.7 were determined by calculating the ratio

$$\frac{\text{distance migrated from the origin by the product}}{\text{distance migrated from the origin by xylene cyanol FF dye}}$$

[b] In order to resolve larger, slowly migrating T_1 RNase and P-RNase secondary products by DEAE paper electrophoresis at pH 1.7, electrophoresis was carried out for 6–7 hr at 1000 V. Electrophoresis at this voltage and for this length of time causes the mononucleotide products C and U to run off the end of the paper, and consequently their presence could not be detected.

DISCUSSION

The 5′-terminal sequence has been determined for *gal* mRNA made in a purified *in vitro* transcription system and isolated by RNA·DNA hybridization. By utilizing low substrate concentrations and low temperature in a kinetically controlled synthetic reaction, we could prepare *gal* RNA of variable but reproducible size.

Sequences for most oligonucleotides from T_1 and pancreatic RNase digests were directly determined by standard procedures. These oligonucleotides were arranged in a unique order by analysis of the partial digestion fragments produced by carboxymethylated pancreatic RNase. The sequence of the first 77 nucleotides accounts for all major T_1 and pancreatic oligonucleotides from *gal* RNA transcribed in 40 sec under our conditions. A single 5′-terminal T_1 oligonucleotide (T33) was identified, but minor alternate initiations may have been overlooked due to streaking of triphosphate-containing oligonucleotides. Thus any transcription of bacterial DNA other than *gal* carried by the template, λp*gal*8, was negligible under the conditions used.

The amino terminal sequence of UDPgalactose 4-epimerase (the protein specified by the first structural gene) was subsequently determined (D. H. Schlesinger, H. D. Niall, and D. Wilson, manuscript in preparation) and correlates exactly with the RNA sequence of residues 27–77 (Fig. 3). This provides a 5′-terminal sequence of a bacterial mRNA initiated at a wild-type promoter and extending into the first structural gene.

Although the amino terminus of epimerase coincides with the AUG at positions 27–29 of the RNA, the latter codon is

preceded by another AUG (residues 15–17) which is in the same reading frame. Since no nonsense codons are present in the intervening sequence, either AUG might serve as the actual translational initiator codon. A similar situation exists in the Qβ phage replicase initiation sequence AUG AAA UGC AUG (24, 25) where the second AUG is known to be the codon for *N*-formylmethionine (26, 27). For the latter case and most cistrons studied (28, 29) both the formyl group and the terminal methionine are cleaved to leave the penultimate amino acid as the amino terminus of the mature protein. In contrast, if epimerase initiates at the second AUG (residues 27–29) of *gal* mRNA, only the formyl group could be removed. Alternately, if epimerase initiates at the first AUG (positions 15–17), a tetrapeptide Met-Glu-Arg-Ile as well as the formyl group must be removed. The discriminatory removal of formyl, formyl plus methionine, or formyl plus several amino acids is likely determined by the amino-acid sequence or conformation of the nascent polypeptide.

At most, the nontranslated "leader" sequence of *gal* mRNA is 26 nucleotides—somewhat shorter than the 38 observed for *lac* mRNA initiated from a mutant promoter (29) and considerably less than the approximately 150 reported for *trp* mRNA (30). Presumably the 5′-terminal sequence of *gal* mRNA is important for ribosome recognition, yet it is most striking in its lack of similarity to sequences preceding the AUG initiator codons reported for *lac* (29), *trp* (30), and the RNA phage cistrons (28). The sequences U-U-U-G-A and A-G-G-A associated with ribosome binding sites (31) are not present in *gal* mRNA. Also absent from *gal* mRNA is either a nonsense codon in-frame and preceding the AUG codon or an AUG in the loop of an obvious hairpin structure. One feature common to *gal* and several other sequences containing ribosome binding sites is a purine-rich stretch at a short distance preceding the initiation codon. The purine-rich stretch in *gal* includes the sequence G-G-A-G (residues 17–20) separated from the AUG at positions 27–29 by seven residues. This is consistent with the proposal that ribosome binding sites contain part of the sequence G-G-A-G-G-U which can be recognized by the 3′ end of 16S rRNA and which precedes the AUG initiator codon by about this distance (32).

A self-complementary sequence occurs in *gal* mRNA between nucleotides 4 and 18. Twelve of these 15 residues can be base paired (boxed in Fig. 3); however, this sequence could reflect an even larger region in the DNA with partial 2-fold rotational symmetry. This type of symmetry (but a dissimilar sequence) is found in the *lac* operator region (33) and appears at the 5′ terminus of *lac* mRNA initiated from the UV5 mutant promoter (29). Whether this region of *gal* RNA has any functional importance for transcription or translation remains to be studied.

In general the 5′ sequence of *gal* mRNA appears unrelated to the "leader" sequences of *lac* UV5 mRNA or the known part of *trp* mRNA. However, a segment of the *gal* sequence, G-A-A-U-U-A-U-G-A-G-A-G (residues 22–23 underlined in Fig. 3), differs at only two positions from a segment of the *lac* sequence, G-A-A-U-U-G-U-G-A-G-C-G (29, 33). The latter sequence composes the promoter-proximal half of a sequence with 2-fold rotational symmetry in the *lac* operator (33) and includes the site of initiation of transcription from the *lac* UV5 mutant promoter (29).

It is possible that part of the nontranslated terminal sequence of *gal* mRNA corresponds to part of the *gal* operator. Also, it is of interest to determine the sequence of the entire *gal* regulatory region upstream from the initiation site for transcription.

We thank Drs. Sankar Adhya, Max Gottesman, and Donald Court for valuable discussions. Excellent technical assistance was provided by Ms. Su-Fang Su, Frances DeNoto, Carol Anderson, and Lucy Price. The work was supported by grants from the National Cancer Institute and the American Cancer Society. J.S. is the recipient of a Medical Scientists Scholarship Fund award from the Connecticut Mutual Life Insurance Co.

1. Echols, H., Reznichek, J. & Adhya, S. (1963) *Proc. Nat. Acad. Sci. USA* **50**, 286–293.
2. Buttin, G. (1963) *J. Mol. Biol.* **7**, 183–205.
3. Shapiro, J. A. & Adhya, S. L. (1969) *Genetics* **62**, 249–264.
4. Buttin, G. (1963) *J. Mol. Biol.* **7**, 164–182.
5. Kalckar, H. M., Kurahashi, K. & Jordon, E. (1959) *Proc. Nat. Acad. Sci. USA* **45**, 1776–1786.
6. Adhya, S. & Echols, H. (1966) *J. Bacteriol.* **92**, 601–608.
7. Saedler, H., Gullon, A., Feithen, L. & Starlinger, P. (1968) *Mol. Gen. Genet.*, **102**, 79–88.
8. Nakanishi, S., Adhya, S., Gottesman, M. E. & Pastan, I. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 334–338.
9. Nakanishi, S., Adhya, S., Gottesman, M. E. & Pastan, I. (1973) *J. Biol. Chem.* **248**, 5937–5942.
10. Miller, Z., Varmus, H. E., Parks, J. S., Perlman, R. L. & Pastan, I. (1971) *J. Biol. Chem.* **246**, 2898–2903.
11. Parks, J. S., Gottesman, M., Perlman, R. L. & Pastan, I. (1971) *J. Biol. Chem.* **246**, 2419–2424.
12. Nissley, S. P., Anderson, W. B., Gottesman, M. E., Perlman, R. L. & Pastan, I. (1971) *J. Biol. Chem.* **246**, 4671–4678.
13. Hua, S. & Markovitz, A. (1974) *Proc. Nat. Acad. Sci. USA* **71**, 507–511.
14. Anderson, W. B., Schneider, A. B., Emmer, M., Perlman, R. L. & Pastan, I. (1971) *J. Biol. Chem.* **246**, 5929–5937.
15. Berg, D., Barrett, K. & Chamberlin, M. (1971) in *Methods in Enzymology*, eds. Grossman, L. & Moldave, K. (Academic Press, New York and London), Vol. 21, pp. 506–519.
16. Heinrikson, R. L. (1966) *J. Biol. Chem.* **241**, 1393–1405.
17. Blattner, F. F. & Dahlberg, J. E. (1972) *Nature New Biol.* **237**, 227–232.
18. Gillespie, D. & Spiegelman, S. (1965) *J. Mol. Biol.* **12**, 829–842.
19. Bøvre, K., Lozeron, H. A. & Szybalski, W. (1971) *Methods Virol.* **5**, 271–292.
20. Sanger, F., Brownlee, G. G. & Barrell, B. G. (1965) *J. Mol. Biol.* **13**, 373–398.
21. Lebowitz, P., Radding, C. & Weissman, S. M. (1971) *J. Biol. Chem.* **246**, 5120–5139.
22. Lee, J. C., Ho, N. W. Y. & Gilham, P. T. (1965) *Biochim. Biophys. Acta* **95**, 503–504.
23. Contreras, R. & Fiers, W. (1971) *FEBS Lett.* **6**, 281–283.
24. Staples, D. H. & Hindley, J. (1971) *Nature New Biol.* **234**, 211–212.
25. Steitz, J. A. (1972) *Nature New Biol.* **236**, 71–75.
26. Weiner, A. M. & Weber, K. (1971) *Nature* **234**, 206–209.
27. Skogeroon, L., Roufa, D. & Leder, P. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 276–279.
28. Weissmann, C., Billeter, M. A., Goodman, H. M., Hindley, J. & Weber, H. (1973) *Annu. Rev. Biochem.* **42**, 303–328.
29. Maizels, N. M. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 3585–3589.
30. Bronson, M. J., Squires, C. & Yanofsky, C. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 2335–2339.
31. Robertson, H. D., Barrell, B. G., Weith, H. L. & Donelson, J. E. (1973) *Nature New Biol.* **241**, 38–40.
32. Shine, J. & Dalgarno, L. (1974) *Proc. Nat. Acad. Sci. USA* **71**, 1342–1346.
33. Gilbert, W. & Maxam, A. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 3581–3584.