Volume 7 Number 7 1979 Nucleic Acids Research

New genes and promoters suggested by the DNA sequence near the end of the coliphage T7 early operon

John C.Boothroyd and Richard S.Hayward

Department of Molecular Biology, University of Edinburgh, UK

Received 28 September 1979

ABSTRACT

We have employed the dideoxynucleotide chain-terminating method to determine the nucleotide sequence of T7 DNA between the physical map positions 18.9% and 19.8%. The most striking features of this sequence are two perfect 21-basepair repeats, each of which appears to contain a promoter for late transcription. In each case the promoter sequence incorporates a putative translational terminator on its left (5'-side of the "sense" strand), and overlaps a potential ribosome-binding site on its right. The region probably lies immediately distal to the early operon, and may contain two short, hitherto unreported protein-coding sequences.

INTRODUCTION

During the first few minutes after infection of Escherichia coli by the bacteriophage T7, transcription is dependent on the host's RNA polymerase and is confined to the "early" operon at the "left" end of the genome (1). This operon is defined by a triple promoter near the left end of the DNA, and a transcriptional terminator close to 19% on the physical map (2,3). The existence of minor internal promoters and terminators has also been reported (4-7).

The early operon (1) codes for one essential and several inessential products, the former being a novel and autonomous T7-specific RNA polymerase, responsible for the transcription of all "late" genes (8,9). Among the products inessential for phage growth in normal laboratory strains and conditions is the T7 DNA ligase, molecular weight about 40,000. This is encoded by gene 1.3, which ends close to the transcriptional terminator of the early operon (10). The first known gene distal to this terminator, gene 1.7, is also inessential, and encodes a protein of unknown function, molecular weight about 17,000 (1).

Late gene transcription is initiated at a large number of promoters specific for T7 RNA polymerase, which fall into two classes: classII

© Information Retrieval Limited 1 Falconberg Court London W1V 5FG England

promoters are active only during the middle period of the infectious cycle, while class III function until the time of lysis (11,12). Class III promoters lie in the right half of the genome, and class II to the left. The leftmost class II promoter is located near 14.6%, and its DNA sequence has been determined (13). Recently another such promoter has been mapped near the terminator of early transcription (14). Accordingly several sites of functional interest should be present near map position 19%. Here we report a 370-basepair nucleotide sequence from this region.

MATERIALS AND METHODS

E.coli B and coliphages $T7^+$, T7 LG37 and T7 C5LG3 am342a were obtained from F.W. Studier. The latter phage was grown on E.coli BBW/1 (Sup^+) provided by R. Hausmann.

Restriction endonucleases HpaI and HpaII were prepared as by Sharp et al (15); HaeII, HinII and some HpaI were from New England Biolabs; HaeIII and HhaI were donated by A. Smith and K. O'Hare, respectively. DNA polymerase I (Klenow A fragment) was from Boehringer Corp. Agarose was from Miles or Sigma (type II), and acrylamide gel components were from Serva. Unlabelled nucleotides were from P-L Biochemicals; poly r(U,G) was from Miles; and [α-^{32}P]dGTP was from the Radiochemical Centre.

Phage growth, DNA purification, and separation of DNA strands were as described previously (16). The reaction conditions for digestion of DNA by restriction endonucleases were: for HaeII and HinII, as specified by New England Biolabs; for HaeIII, 6.6 mM Tris. HCl, pH 7.4, 1 mM DTT, 6.6 mM $MgCl_2$; for HhaI, 6 mM Tris. HCl, pH 7.9, 6 mM 2-mercaptoethanol, 6 mM $MgCl_2$, 100 mM NaCl; and for HpaI, 10 mM Tris. HCl, pH 7.5, 10 mM 2-mercaptoethanol, 10 mM $MgCl_2$, and 80 mM NaCl. Incubation was typically 1-3 h at 37°. Horizontal agarose, and vertical polyacrylamide gel electrophoretic fractionation of DNA digests were performed essentially as by McDonell et al (2) and Maniatis et al (17) respectively. DNA in gels was located by staining with ethidium bromide, and photographing under UV illumination. Agarose gels were stained for 20 min in 1 µg dye/ml, then rinsed in H_2O for 1 h. Polyacrylamide gels were stained for 10 min in 3 µg dye/ml, and rinsed for 8 min.

DNA fragments were recovered from gels by either of two methods. Passive elution (communicated by K. O'Hare) was slower than electro-elution, but gave a similar recovery (50-100%), and a better preparation. The required regions of preparative gels were first identified by excising thin longitudinal slices,

staining, and photographing as above. Using the photograph as a guide, the appropriate region of the unstained remainder of the gel was excised, and crushed by passage through a sterile plastic syringe. For passive elution, the gel pieces were mixed with 5 ml of PE buffer (50 mM Tris. HCl, pH 8.0, 100 mM NaCl, 10 mM EDTA) in a siliconised Corex tube, sealed with Parafilm, and swirled gently overnight, 37°. The supernatant was recovered by centrifugation (5000 g, 5 min), and combined with a PE wash (3 ml) of the loose pellet. It was then passed through a 1 ml DEAE-cellulose column (Whatman DE-52, PE buffer, in a siliconised Pasteur pipette plugged with glass wool). The flow-through was recycled, and the column then washed with PE buffer (3 ml). The DNA was eluted in PE buffer (1-1.5 ml) containing 1 M NaCl, and precipitated by the addition of ethanol (2 vols). If the product was to be used in sequencing experiments, single-stranded DNA template was added prior to the ethanol.

Electro-elution from the crushed gel was as described by McDonell *et al* (2), except that the final contents of each sac were filtered through a Whatman GF/C filter (1-2 cm Hg negative pressure) prior to ethanol precipitation.

DNA sequencing was carried out as described by Sanger *et al* (18), with the following modifications. (*i*) To minimise priming by the 3' ends of the linear template DNA strands, these ends were first "blocked" by dideoxy-ribothymidine addition. The DNA (at 30 µg/ml) was denatured in HN buffer (6.6 mM Tris. HCl, pH 7.4, 6.6 mM $MgCl_2$, 1 mM DTT, 50 mM NaCl), 100°, 3 min. Intrastrand annealing was encouraged by incubating in a water bath cooling from 50° to 20° over a 2 h period. 0.03 mM dATP, dGTP and dCTP, 0.25 mM 2',3'-dideoxy TTP, and *E.coli* dpolI (Klenow A fragment: final concentration 5 units/ml) were then added, and the solution incubated for 1 h, 20°. Polymerisation was stopped by addition of EDTA to 20 mM, and heating at 70°, 10 min. The mixture was cooled, passed through Sephadex G-100 (13 x 0.8 cm diam, equilibrated with 5 mM Tris-acetate, pH 8.0, 0.25 mM EDTA, 20°), and eluted with the same buffer, flow-rate 10 ml/h; 0.5 ml fractions were collected. The first UV-absorbing peak, containing only polynucleotide, was pooled and stored at 2°. (*ii*) Template and primer were annealed for 1 h (rather than 15 min) at 60°. (*iii*) The relative concentration of dideoxy NTP was 2-fold higher than that stipulated by Sanger *et al* (18). (*iv*) The sequencing gels were finally fixed in 10% (v/v) acetic acid for 10 min, washed in H_2O for 2 min, and dried onto Whatman 3MM paper using a Bio-rad gel drier (20 min). They were then taped to a glass plate, and overlaid

with light-activated X-ray film (pre-exposed to a single flash from 50 cm by an Agfatronic 150A flash unit, through an Ilford 6B filter). An Ilford Fast Tungstate or Dupont Cronex intensifying screen was then laid over the film, followed by a second glass plate. The whole sandwich was clamped, wrapped in Al foil and black polythene, and exposed at -70°.

RESULTS

Restriction mapping. To determine which fragments of DNA would be required as primers in the chain-terminating approach to DNA sequencing, it was necessary to establish a map of restriction endonuclease cutting sites in the region of interest. For this purpose we exploited the deletion LG37, which eliminates the transcriptional terminator and whose endpoints have been mapped to 14.5 (±0.2)% and 19.4 (±0.1)% (Fig. 1 and refs. 10,19,20) and the deletion LG3. ΔLG3 does not affect the terminator; its left limit appears identical to that of ΔLG37, while its rightward endpoint has been mapped at

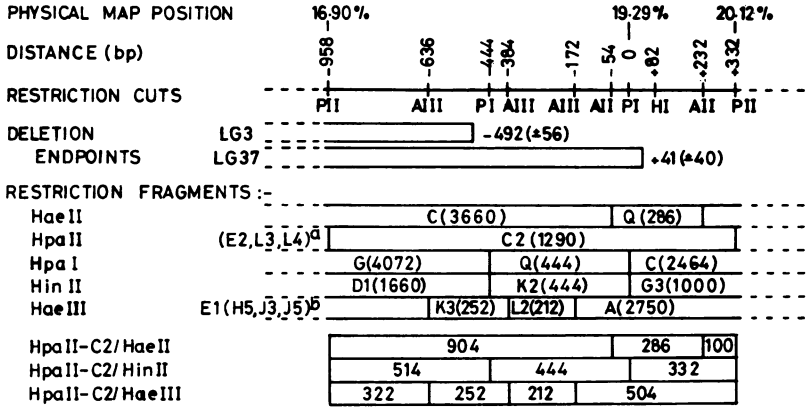

Fig. 1 Cleavage map of the region of coliphage T7 DNA between physical map positions 16.9 and 20.1%. Endonuclease cuts are shown as PI (HpaI and HinII), PII (HpaII), AII (HaeII), AIII (HaeIII), and HI (HhaI). The HhaI site at +82 is predicted by restriction analyses of Studier et al (20) and Boothroyd (19). The positions of the cuts are given in basepairs relative to the HpaI site mapped at 19.29% by McDonell et al (2). The estimates of fragment sizes are also given in bp. The HpaII-C2 subfragment sizes have been normalised to total 1290 bp: this assumes that there were no undetected small fragments.
a: HpaII fragments E2, L3 and L4 are 760, 135 and 125 bp, respectively.
b: HaeIII fragments E1, H5, J3 and J5 are 1140, 550, 345 and 330 bp, respectively.
Our estimates of the rightward endpoints of deletions LG3 and LG37 are also shown.

18.1 (±0.2)% (ibid.). The strain of T7 used to study the effects of ΔLG3 also contained the deletion C5 and the amber mutation 342a (10), but this did not substantially complicate our analysis. Our estimates of the lengths of the deletions, based on restriction analysis, are: LG37, 1946 bp; LG3, 1426 bp; C5, 1844 bp.

Since the approach adopted to derive our cleavage map was in no way unusual, we shall not present the data or interpretation here. A supplement containing these is available on request (to R.S.H.). Our conclusions are presented in Fig. 1. They are in full agreement with the independent analyses of Studier et al (20), and confirmed by the nucleotide sequence results reported below. For simplicity we have adopted the nomenclature of Studier's group for the fragments produced by restriction endonucleases. The estimates of fragment lengths given in Fig. 1 are closely based on those of Studier et al (20), and we have confirmed them by comparison with restriction fragments of plasmid pBR322, whose nucleotide sequence is known (Sutcliffe, pers. comm.).

Sequence determination. Fig. 2 shows the primers used and the regions for which sequence information was obtained, in ten independent experiments. As an example of our raw data, an autoradiograph derived from experiment h of

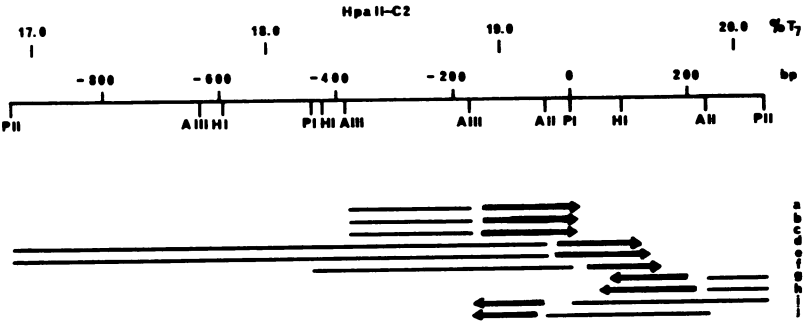

Fig. 2 Summary of sequencing experiments. The figure includes a detailed map of the cutting sites for various restriction endonucleases (determined and represented as in Fig. 1) within the HpaII fragment C2. Note that all PI (HpaI) sites are also recognised by HinII, and all AII sites by HhaI. The thin lines beneath the map represent the various primers used, with the thicker arrows showing the direction of extension and (by their positions) the region for which useful sequence data were obtained. In experiments a-f, the template was $T7^{+}$ r-strand: in the remainder it was l-strand. The primers used were: a-c, HaeIII-L2; d,e, HpaII-C2/HaeII-904 bp (the 904 bp fragment produced by recutting HpaII-C2 with HaeII); f, HinII-K1; g,h, HpaII-C2/HaeII-100 bp; i, HpaIIC2/HinII-332 bp; j, HaeII-Q.

Fig. 2 is shown in Fig. 3. All results have been collated to produce the final sequence presented in Fig. 4. This spans from -163 to +207 nucleotides relative to the HpaI cut which separates HpaI fragments Q and C, and is located at 19.29 on the physical map of T7. As shown in Fig. 4, nine nucleotides could not be assigned unambiguously. Eight of those ambiguities lie within the HpaI-Q portion of our sequence, which has been independently determined by J. Dunn (pers. comm.) using the method of Maxam and Gilbert (21). Dunn's sequence is almost completely in agreement with ours. In the five positions between -149 and -132 where we have indicated alternative

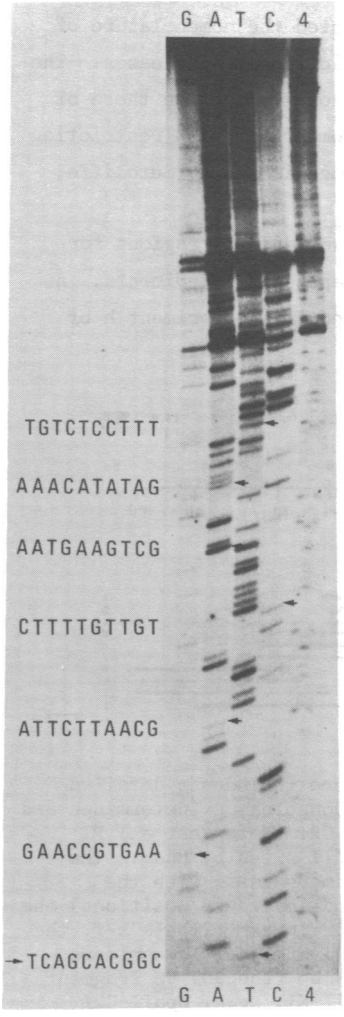

Fig. 3. Representative sequencing-gel autoradiograph. Chain terminating method of Sanger et al (18). Primer used: $T7^{+}$-HpaII-C2/HaeII - 100 bp. Template: $T7^{+}$ l-strand DNA. The labelled nucleotide was [α-^{32}P] dGTP, and the products were digested with HaeII prior to gel analysis. In the track marked 4, a small sample from each of the individual reactions (i.e. G, A, T and C) was loaded to provide a reference ladder of bands. The gel was run for 2 h at 25 mA. The film was exposed for 15 h. To aid interpretation, every tenth band has been marked ←, and the derived sequence +207 to +138 (complementary to that given in Fig. 4) is shown on the left. The band occurring in all four tracks at a position corresponding to nucleotide +127 is about the right size to be labelled primer (232 - 127 = 105 bp: these extensions were to the left from the HaeII site at +232, so that a band at +127 represents a fragment about 105 nucleotides in length). The origin of the similar band at position +101 is unknown.

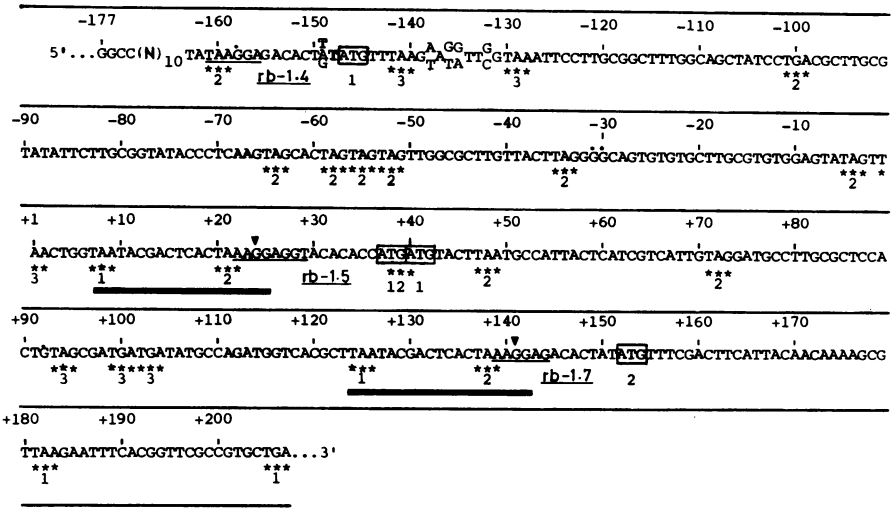

Fig. 4. Nucleotide sequence and proposed functional sites between map positions 18.9 and 19.8. Only the 1- or sense-strand of DNA is shown: nucleotides are numbered from the HpaI site at 19.29 (ref. 2). Putative functional sites (see text) are marked.

Nucleotide designations: N, unidentified; A/G, alternative possibilities; Ġ, tentative or probable; G, considered certain.

Putative translational features: TGA (3), stop-codon and phase; ATG (1), start-codon and phase; ..AGGA.., region of complementarity to 3'-terminus of E.coli 16S rRNA; rb-1.4, name of proposed ribosome-binding site.

Putative transcriptional features: ..ATA.., class II promoter; Ġ, class II RNA start-site.

nucleotides, Dunn's results unambiguously confirm our uppermost alternative. Dunn assigns C instead of G at position -32, but this, the only discrepancy, does not affect the discussion of our results.

Structural features of the sequence. A computer-assisted search predicted only two possibilities for the formation of stable stem-loop structures in an RNA transcript of the region sequenced. These are diagrammed in Fig. 5. We have no evidence for their physical reality.

There are several homologous regions within the sequence. Aside from the 21 bp direct repeats at +7 to +27 and +124 to +144, discussed below, the region -160 through -142 has 18/19 nucleotides which are identical to those at +139 through +157. This homology probably has a functional significance (both regions are candidate translational initiators) and could also reflect an evolutionary relationship. We also note a nearly tandem duplication at -96 to -87 and -84 to -74 (exact if the G at -78 is looped out), while a closely related sequence appears at -20 through -11.

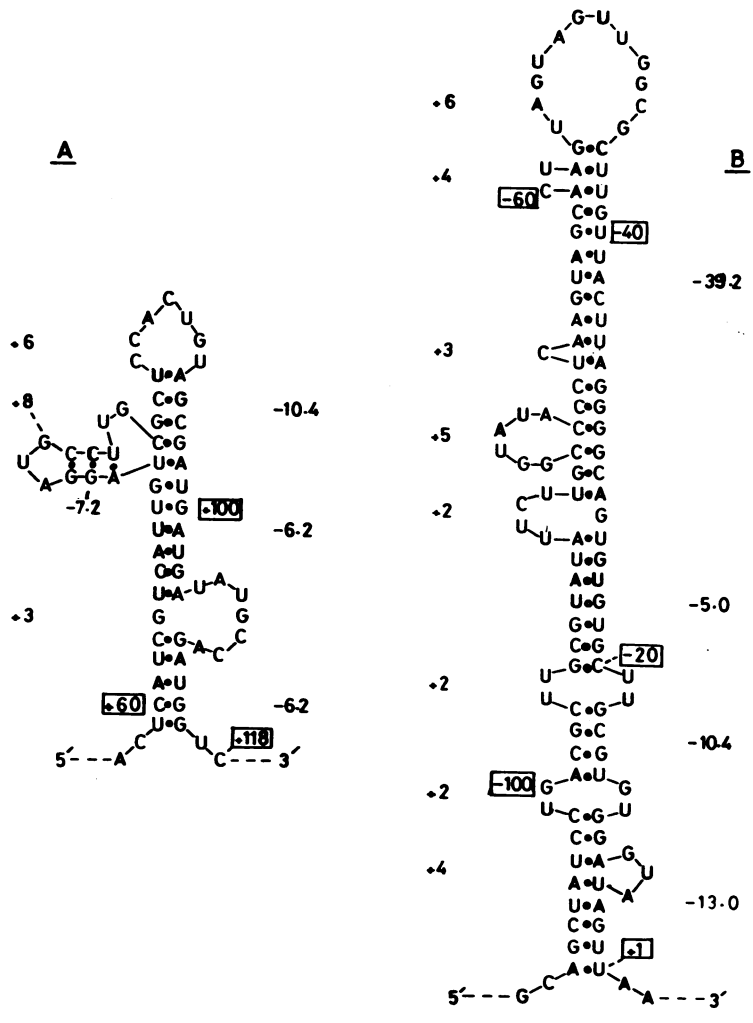

Fig. 5. Most stable predicted RNA secondary structures.

a. Possible base-pairing in an RNA transcript containing the region from +59 through +116 (numbering of nucleotides (in boxes) is as in Fig. 4). The predicted free energy change of formation for this structure (22) is -13 kcal/mol. The contribution of each segment to this overall value is given alongside the drawing.

b. Possible base-pairing in an RNA transcript containing the region from -109 through -1. The predicted free energy change of formation for this structure is -39.6 kcal/mol.

The following restriction endonuclease recognition sites, predicted by the sequencing experiments, are in good agreement with restriction analyses (Fig. 1; and ref. 20): HaeIII (5'-nucleotide of sequence is at -177), HaeII (-48), HpaI (-3), HhaI (+82). Several unconfirmed sites are also predicted: EcoRI* (-128 and +185), BbSl (-111), AluI (-109), MnlI (-72 and +25), HinfI (+13 and +130) and TaqI (+157): see Roberts (23) for the relevant sequences.

DISCUSSION

Possible promoters for late transcription. The most striking feature of the DNA sequence which we have determined is the presence of two perfect 21 bp direct repeats. As shown in Fig. 6 the first 19 nucleotides of this repeated sequence are identical (except at position 16) with the sequences of four class III promoters for late transcription, determined by Rosa (24) and Oakley et al (25), and show almost as much homology with the class II late promoter located at 14.6% (13). Recently Panayotatos and Wells (26) have reported two very similar class II promoter sequences, located at 14.8 and 15.9%. Note that there are two closely spaced promoters at 14.6 and 14.8%. Although we have no direct evidence, it seems very likely that our repeated

			T7 Map Position	Class
5'...GAAAT	TAATACGACTCACTATAGGGAGA	CCACA...3'	57	III
GGAAA	TAATACGACTCACTATAGGGAGA	GGCGA	87	III
GAATT	TAATACGACTCACTATAGGGAGA	CCTCA	55	III
TAAAT	TAATACGACTCACTATAGGGAGA	TAGGG	46.5	III
CAAAT	CAATACGACTCACTATAGAGGGA	CAAAC	14.6	II
+2 ACTGG	TAATACGACTCACTAAAGGAGGT	+34 ACACA	19.35	II?
+119 ACGCT	TAATACGACTCACTAAAGGAGAC	+151 ACTAT	19.65	II?

Fig. 6. Late promoters. The nucleotide sequences of one class II (13) and four class III (24,25) T7 Rpol promoter regions are compared with those of the two putative class II promoters identified in the present work. (The numbering of the nucleotides in the latter two cases is as in Fig. 4.) Only the 1- or sense-strand sequence is shown. Approximate locations on the physical map of T7, and the class of the promoter, are given on the right. The RNA start-site, where known, is indicated by * above the appropriate nucleotide. The sequences are aligned by homology with the 23 base-pairs (enclosed within the box) which are identical in the four class III promoters. Homologous nucleotides within this box are underlined.

sequences (closely spaced near 19.35 and 19.65%) function as class II promoters for T7 RNA polymerase. Presumably they correspond to the promoter mapped at approximately 18.9% by Kassavetis and Chamberlin (14). The remaining sequence of the HpaI-Q fragment of T7 (limits 18.2-18.9%, leftward of the region which we have studied) contains no likely promoter sequences for the T7 polymerase (Dunn, pers. comm.).

Translational features. We consider it unlikely that the sequence presented in Fig. 4 contains any frameshift errors, and assume none throughout the following discussion.

The RNA transcript corresponding to this sequence would contain three potential ribosome-binding and translation-initiation sites, which we have tentatively named rb-1.4, rb-1.5, and rb-1.7. As shown in Fig. 7 these ribosome-binding sites are distinguished by sequences, 6-8 nucleotides in length, which are perfectly complementary to the 3'-terminal sequence of E.coli 16S ribosomal RNA. Such "Shine and Dalgarno" complementarities are well-established features of E.coli ribosome-binding sites on mRNA (27,28). Fig. 7 also shows that each of our "Shine and Dalgarno" sequences is followed by an AUG start codon (7 nucleotides downstream), and flanked by sequences displaying additional homologies with the known initiator of T7 gene 0.3 translation (rb-0.3; ref. 29), as well as with the putative initiator for gene 1 (rb-1; ref. 30). These particular additional homologies are not observed in the majority of known

```
5'.......AACUGCACGAGGUAACACAAGAUG...3'        rb-0.3
          *****      ...

   GAUUUACUAACUGGAAGAGGCACUAAAUG              rb-1
            ****      ...

        -163              -145
          UAUAAGGAGACACUUUAUG                  rb-1.4
         *******      ...

  +11                       +39
   ACGACUCACUAAAGGAGGUACACACCAUG               rb-1.5
        ********      ...

 +126                       +154
   AUACGACUCACUAAAGGAGACACUAUAUG               rb-1.7
        ******      ...
```

Fig. 7. Ribosome-binding sites. The nucleotide sequences of the gene 0.3 ribosome binding site (rb-0.3; ref. 29) and of the proposed gene 1 site (rb-1; ref. 30) are compared with the three putative sites suggested for genes 1.4 (rb-1.4), 1.5 (rb-1.5) and 1.7 (rb-1.7). (The numbering of nucleotides in the latter three cases is as in Fig. 4.). The sequences are aligned by the AUG codon, which is indicated by subscript dots. Subscript asterisks indicate nucleotides complementary to the 3'-terminal region of E.coli 16S RNA. Further regions of homology between rb-1.7 and the other sequences are underlined.

binding sites for E.coli ribosomes. Such features may be important for ribosome-binding during the course of T7 infection, and/or could reflect some common evolutionary origin. The region which we have sequenced is certainly transcribed, at least in the late period after infection (for example from the class II promoter at 14.6%; ref. 12). We consider it likely that the putative ribosome-binding sites function for translation in vivo, although direct evidence is clearly desirable (cf. ref. 31).

rb-1.4 is followed (Fig. 4) by 50 sense codons in phase with its AUG start codon, and then by a UAA at +8. Since gene 1.3 probably terminates some 40 nucleotides to the left of rb-1.4 (J. Dunn and F.W. Studier, pers. comm.) we suggest that a previously unreported gene may be located between 18.92% and 19.30% on the physical map of T7, and that it should be called gene 1.4, in line with the standard system of T7 gene nomenclature (1).

The RNA of rb-1.5 includes tandem AUG codons, followed in phase by 27 sense codons and than a UAA stop codon (at +125). rb-1.5 is preceded by stop codons in all three phases. Thus we suggest that there may be another previously unreported protein-coding region, gene 1.5, initiated from rb-1.5 and mapping at 19.38-19.60%.

The RNA of rb-1.7 includes an AUG at position +152 (19.67%), followed in phase by at least 19 sense codons up to the end of the known sequence. This could be the beginning of the known late gene 1.7. Further work will be necessary to confirm this suggestion, and to test for the putative products of genes 1.4 and 1.5. It should be noted that the products of short genes might well have gone undetected by the methods so far applied to the molecular genetics of T7. Three additional short genes have only recently been discovered in the well-studied early operon (20).

Each of our putative promoter sequences for late transcription contains a UAA triplet at the 5'side, which probably terminates a protein-coding unit in the RNA transcript, as described above. The repeated sequences also include a run of purines at their 3'-ends, which not only contain likely start sites for late transcription, but also the "Shine and Dalgarno" components of putative translation-initiation sites. We are struck by the frugality of this arrangement, which would maximise the use of intercistronic sequences, and may conceivably have played a role in the evolution of the promoter sequence for T7 RNA polymerase.

Terminator of the early operon. The terminator for early transcription maps near the HpaI site at position zero in our determined sequence (2). (It is interesting to note that this terminator is not recognised by the T7-coded RNA

polymerase, during late transcription in vivo (12) or in vitro (11).) Previous work from our laboratory suggested that the primary RNA product of the early operon terminates with a 3'-C residue, within a sequence 5'-PyCG(A,T or C)-3' on the sense strand of the DNA (16,32,33). This result was based on analysis of the products of reactions in which an in vitro transcript of T7 DNA was hybridised to the complementary DNA strand, and its 3'-terminus used as a primer for DNA polymerase in the presence of [α-^{32}P] dNTP. Independent application of the periodate [^{3}H]$NaBH_4$ method for RNA-3'-terminal sequencing suggested that the final three bases of the RNA are C (34). However, the sequence CCC occurs nowhere in the currently determined region of the DNA. Note that runs of T, runs of A, and the sequence CAATCAA, which have been observed at various transcriptional terminators (35-38), are also absent from this region of T7 DNA.

The sequence PyCG(A,T or C) occurs at positions +63 and +158. It may be noted that neither position is associated with any significant two-fold rotational symmetry in the DNA sequence. Since ΔLG37 inactivates the normal terminator of the early operon, and since its rightward endpoint is to the left of nucleotide +82, it seems unlikely that the terminator is located as far to the right as +159. Setting aside our periodate-based sequence results, position +64 is a possible candidate for the transcriptional stop of the early operon. Alternatively the stop site could lie to the left of the sequence which we have studied. However, we believe that it should lie to the right of nucleotide -175 (which according to our preliminary sequence data is the position of the restriction endonuclease cut separating HaeIII fragments L2 and A), since we have evidence that RNA polymerase bound at the terminator lies within HaeIII fragment A (O'Hare, Boothroyd and Hayward, in preparation). Recently we have learned that independent DNA and RNA sequence determination (Dunn, pers. comm.) strongly suggest that the early RNA terminates at nucleotides -166 and -167.

We thank Drs N. Brown, J. Dunn, D. McConnell, K. and N. Murray, K. O'Hare, F. Sanger, A. Smith and F.W. Studier and Mr A. Newman for discussion, exchange of unpublished information and gifts or exchange of material; Dr J. Collins for the computer study; and the Medical Research Council for Project Grant Support. J.C.B. was an Overseas Science Research Scholar of the Royal Commission for the Exhibition of 1851.

Address requests for offprints to R.S.H.

New address for J.C.B.: Department of Immunochemistry, Wellcome Research Labs. Langley Court, Beckenham, Kent, UK

REFERENCES

1. Studier, F.W. (1972) Science 176, 367-376
2. McDonell, M.W., Simon, M.N. and Studier, F.W. (1977) J.Mol.Biol. 110, 119-146
3. Siebenlist, U. (1979) Nucl.Acids Res. 6, 1895-1907
4. Delius, H., Westphal, H. and Axelrod, N. (1973) J.Mol.Biol. 74, 677-687
5. Minkley, E. and Pribnow, D. (1973) J.Mol.Biol. 77, 255-277
6. Stahl, S.J. and Chamberlin, M.J. (1977) J.Mol.Biol. 112, 577-601
7. McConnell, D. (1979) Nucl.Acids Res. 6, 525-544
8. Chamberlin, M.J., McGrath, J. and Weskell, L. (1970) Nature 228, 227-231
9. Chamberlin, M.J. and Ring, J. (1973) J.Biol.Chem. 248, 2235-2244
10. Simon, M.N. and Studier, F.W. (1973) J.Mol.Biol. 79, 249-265
11. Golomb, M. and Chamberlin, M.J. (1974) Proc.Nat.Acad.Sci.USA 71, 760-764
12. Pachl, C.A. and Young, E.T. (1978) J.Mol.Biol. 122, 69-101
13. Oakley, J.L. and Coleman, J.E. (1977) Proc.Nat.Acad.Sci.USA 74, 4266-4270
14. Kassavetis, G.A. and Chamberlin, M.J. (1979) J.Virol. 29, 196-208
15. Sharp, P.A., Sugden, B. and Sambrook, J. (1973) Biochem. 12, 3055-3063
16. Peters, G.G. and Hayward, R.S. (1974) Eur.J.Biochem. 48, 199-208
17. Maniatis, T., Jeffrey, A. and van de Sande, H. (1975) Biochem. 14, 3787-3794
18. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc.Nat.Acad.Sci.USA 74, 5463-5467
19. Boothroyd, J.C. (1979) PhD Thesis, University of Edinburgh
20. Studier, F.W., Rosenberg, A.H., Simon, M.N. and Dunn, J.J. (1979) J.Mol.Biol. (in press)
21. Maxam, A. and Gilbert, W. (1977) Proc.Nat.Acad.Sci.USA 74, 560-564
22. Tinoco, I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Nature New Biol. 246, 40-41
23. Roberts, R.J. (1978) Gene 4, 183-193
24. Rosa, M.D. (1979) Cell 16, 815-825
25. Oakley, J.L., Strothkamp, R.E., Sarris, A.H. and Coleman, J.E. (1979) Biochem. 18, 528-537
26. Panayotatos, N. and Wells, R.D. (1979) Nature 280, 35-39
27. Shine, J. and Dalgarno, L. (1974) Proc.Nat.Acad.Sci.USA 71, 1342-1346
28. Steitz, J.A. and Jakes, K. (1975) Proc.Nat.Acad.Sci.USA 72, 4734-4738
29. Steitz, J.A. and Bryan, R.A. (1977) J.Mol.Biol. 114, 527-543
30. McConnell, D. (1979) Nucl.Acids Res. (in press)
31. Blumberg, B.M., Nakamoto, T. and Kêzdy, F.J. (1979) Proc.Nat.Acad.Sci.USA 76, 251-255
32. Hayward, R.S. and Peters, G.G. (1974) FEBS Proc. 33, 361-366
33. O'Hare, K.M. (1978) PhD Thesis, University of Edinburgh
34. Peters, G.G. and Hayward, R.S. (1974) Biochem.Biophys.Res.Comm. 61, 809-816
35. Gilbert, W. (1976) in RNA Polymerase, Lodick, R. and Chamberlin, M.J. Eds., pp. 193-205 Cold Spring Harbor Laboratory, New York
36. McMahon, J.E. and Tinoco, I. (Jnr.) (1978) Nature 271, 275-277
37. Rosenberg, M., Court, D., Shimatake, H., Brady, C. and Wulff, D. (1978) Nature 272, 414-423
38. Küpper, H., Sekiya, T., Rosenberg, M., Egan, J. and Landy, A. (1978) Nature 272, 423-428