Transcription antitermination by phage λ gene Q protein requires a DNA segment spanning the RNA start site

Xianjie Yang, Craig M. Hart, Elizabeth J. Grayhack,[1] and Jeffrey W. Roberts

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853 USA

The gene *Q* protein of phage λ is a transcription antiterminator that modifies RNA polymerase near the phage late gene promoter and thereby causes antitermination at distant sites. To define the site of action of Q protein, we have reconstructed the regulatory system on plasmids that allow the intracellular concentration of Q protein to be regulated, and that allow the effect of Q protein on transcription from variant promoter segments to be measured in vivo and in vitro. We show that DNA sequences essential for Q protein-mediated antitermination span the RNA start site, but do not extend beyond nucleotide 18 of the late RNA coding region. We also show that the modification that permits antitermination persists while RNA polymerase passes at least two terminators in vivo and in vitro.

[*Key Words*: Antitermination; bacteriophage λ; Q protein]

Received November 15, 1986; revised version received and accepted January 30, 1987.

Transcription antitermination mechanisms are important to genetic regulation in bacteria. Furthermore, there are recurrent suggestions that termination or elongation of transcription is regulated in eukaryotes as well (for example, see Bentley and Groudine 1986; Gilmour and Lis 1986). The best-characterized transcription antiterminators are the products of the *Escherichia coli* bacteriophage λ genes *N* and *Q* (Friedman and Gottesman 1983). The λ *Q* gene protein activates phage late gene expression by allowing RNA polymerase to pass a transcription terminator, $t_{R'}$, which is located 194 nucleotides downstream from the initiation site of the single late gene promoter $p_{R'}$ (Fig. 1), and thereby to transcribe the 23-kb late gene operon. In vitro, purified λ Q protein acts on RNA polymerase to allow this same antitermination event in a well-defined transcription reaction containing only RNA polymerase, the NusA transcription factor, Q protein, DNA containing the λ late promoter, and small molecules (Grayhack and Roberts 1982; Grayhack et al. 1985).

Q protein antiterminates transcription initiated only at the λ late promoter or the nearly identical *Salmonella* phage P22 late promoter (Grayhack et al. 1985). It does not act even on the late promoter of phage 82, which is controlled in a similar way by an antiterminator of its own (J. Goliger, X. Yang, and J.W. Roberts unpubl.). These facts argue strongly that Q protein recognizes a specific sequence in DNA or RNA, where it engages RNA polymerase and modifies the termination properties of the enzyme. Likewise, the λ gene *N*-encoded antiterminator, which regulates phage early gene expression, recognizes a site named *nut* (for *N* utilization) where it modifies RNA polymerase. However, whereas N protein antiterminates transcription from any promoter as long as *nut* is placed in the transcribed sequence (Salstrom and Szybalski 1978; deCrombrugghe et al. 1979), the site required for Q function (*qut*) has not been separated from the late gene promoter. Deletion analysis has identified sequences within the first 20 nucleotides of the late gene transcript coding region, close to the −35 and −10 elements of the promoter itself, which are essential to Q function (Grayhack and Roberts 1982; Somasekhar and Szybalski 1983; Grayhack et al. 1985). More direct evidence for the involvement of a segment just after the RNA start is the fact that Q protein can modify RNA polymerase if Q protein is added while the enzyme is pausing during transcription in vitro at a natural pausing site at about nucleotide 16 of the late gene transcript (Grayhack et al. 1985; Fig. 1); it seems likely that this pause is essential to Q function. Thus, the nucleotide sequences that allow engagement of Q protein must be very close to, and possibly partly within, the late gene promoter.

To define more precisely the structure on which Q protein acts at the late gene promoter, we have reconstructed this regulatory system on plasmids that allow us to measure the effect of Q protein on transcription initiated at the same promoter segments in vivo and in vitro. By testing modified promoter segments in this system, we show the following: (1) Q function at the λ late promoter requires the natural sequences both before and after the RNA start site, and Q protein does not act

[1]Current address: Department of Biology, University of Rochester, Rochester, New York 14627 USA.

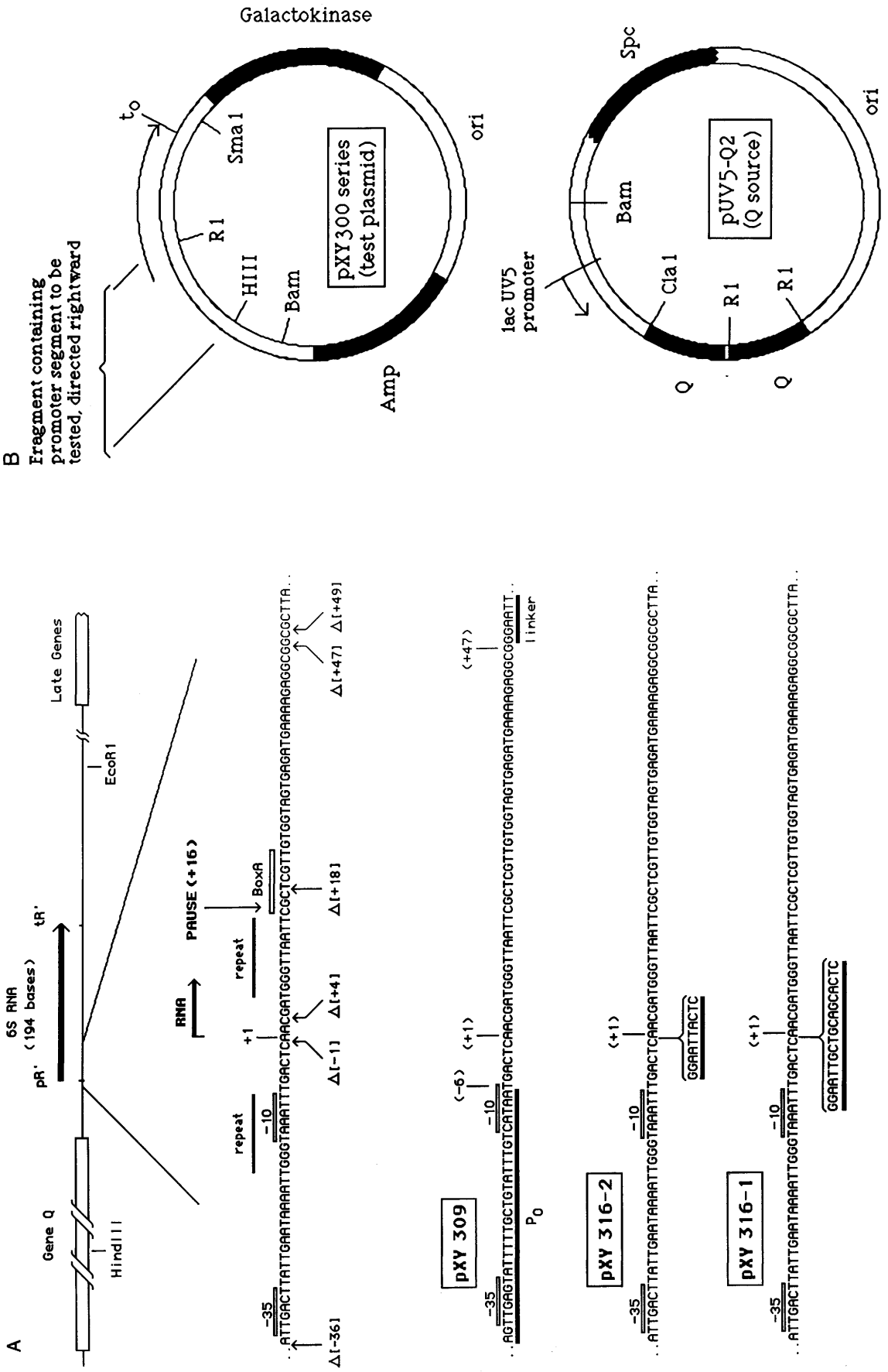

Figure 1. Maps and sequences of DNAs derived from the late gene regulatory region of phage λ. (*a*) The wild-type regulatory region and three variants. Gene *Q* itself is transcribed rightward, from a promoter in the early region of the phage DNA. The *Hin*dIII site within gene *Q* is about 450 bp from the initiation site of the 6S RNA. The *Eco*RI site is 190 bp to the right of t_R. The pause site at +16, the repeats, and the sequence "BoxA" were discussed previously (Grayhack et al. 1985). Vertical arrows and Δ symbols indicate end points of deletions made as described. The −35 and −10 recognition elements of promoter $p_{R'}$ are shown; RNA initiates at nucleotide +1. In the three variants, heavy underlining designates DNA substituted for, or added to, the original sequence. (*b*) Structures of the pXY300 series plasmids and plasmid pUV5-Q2. Amp and Spc designate genes conferring resistance to ampicillin and spectinomycin. The origin of replication is designated ori.

if these sequences are separated by a few nucleotides. (2) The sequence requirements for Q function in vivo and in vitro are identical, a result that confirms the authenticity of the in vitro reaction. (3) Once Q protein-modified RNA polymerase passes the first of two successive terminators, it then can pass the second with very high efficiency both in vivo and in vitro. (4) A higher concentration of Q protein appears to be required in vivo than in vitro, when Q protein is supplied in *trans* by a plasmid in the cell.

Results

Assay system for Q function

To measure the response of promoter segments to Q protein in vivo, we constructed plasmids that allow the easily assayed enzyme galactokinase to be controlled by Q protein (Fig. 1b). Q protein is made from plasmid pUV5-Q2, which contains two copies of gene *Q* transcribed from the lactose operon promoter UV5, so that Q protein is inducible by IPTG. A second plasmid in the same cell, one of the set pXY301–pXY316, contains the promoter segment to be tested, followed by a terminator ($\lambda t_{O'}$ chosen for convenience) and the galactokinase gene. When $p_{R'}$ and the associated *qut* site are present in the test plasmid, Q protein made from pUV5-Q2 upon induction by IPTG allows transcription initiated at $p_{R'}$ to pass the terminator and make galactokinase. (Since uninduced cells make little galactokinase, this scheme avoids the effect of galactokinase expression on plasmid copy number during long-term growth, as we have confirmed directly.) By 90 min after induction, cells containing a Q-responsive plasmid such as pXY306 have up to 20 times the concentration of galactokinase as uninduced cells (Table 1), or cells lacking the Q source plasmid pUV5-Q2 (data not shown). The numbers in Table 1 underestimate the rate of induced galactokinase synthesis by about a factor of 2 for two reasons. First, the concentration of galactokinase would be proportional to its rate of synthesis only after the original cell mass had been diluted to a negligible concentration through growth, whereas in this experiment the cell mass increased only threefold after induction. Second, there is a lag of about 20 min after IPTG is added before galactokinase synthesis starts.

We found the separation of gene *Q* and promoter $p_{R'}$ onto two separate plasmids to be necessary for regulated Q-dependent galactokinase expression in vivo. A similarly constructed plasmid containing both gene *Q* and promoter $p_{R'}$ in their natural *cis* configuration expressed galactokinase constitutively, despite the presence of lactose repressor which should have repressed the *Q* gene; we attribute this to an autocatalytic effect of slight Q synthesis that allows transcription from $p_{R'}$ to circle the plasmid and make more Q protein, thus bypassing repressor control.

Table 1 shows that galactokinase is inducible by IPTG when the test plasmid contains DNA segments including $\lambda p_{R'}$ (e.g., pXY307) or the late promoter of the related phage P22 (pXY313), both of which respond to λ Q protein during phage growth. As expected, galacto-

Table 1. *Activity of Q protein on promoters in vivo and in vitro*

Plasmid			Galactokinase activity in vivo				Readthrough in vitro (%)	
			experiment 1		experiment 2			
name	promoter	terminator	−IPTG	+IPTG	−IPTG	+IPTG	−Q protein	+Q protein
pXY301	$p_{R'}\Delta(-36)$	t_O	0.41	0.42	0.46	0.35	—	—
pXY302	$p_{R'}\Delta(-1)$	t_O	3.2	3.6	3.7	4.4	5.7	4.8
pXY303	$p_{R'}\Delta(+4)$	t_O	2.2	2.1	2.7	3.0	8.9	7.1
pXY304	$p_{R'}\Delta(+18)$	t_O	3.7	37.2	3.7	74.6	6.4	25.0
pXY305	$p_{R'}\Delta(+47)$	t_O	5.1	51.1	5.2	55.7	4.5	44.0
pXY306	$p_{R'}\Delta(+49)$	t_O	3.1	60.5	4.2	77.3	6.2	50.0
pXY307	$p_{R'}\Delta(+63)$	t_O	2.6	67.0	2.7	60.5	9.2	49.0
pXY310	$p_{R'}$	$t_{R'}$	3.7	87.0	4.6	67.7	<1.0	25.0
pXY308	$p_{R'}$	$t_{R'}$ and t_O	0.33	42.0	0.40	46.9	<1.0	36.0
pXY313	P22 late (+48)	t_O	2.1	48.8	2.4	67.8	8.1	50.0
pXY309	$p_O/p_{R'}$	t_O	3.4	3.2	3.4	3.3	<5.0	<5.0
pXY312	82 late (+52)	t_O	0.59	0.79	0.9	1.0	10.0	10.0
pXY314	p_R (λ early)	t_O	9.2	7.7	11.5	9.9	9.0	5.8
pXY316-1	$p_{R'}$ (18 bp insertion)	t_O	4.3	4.5	5.2	6.8	8.4	7.9
pXY316-2	$p_{R'}$ (10 bp insertion)	t_O	3.0	2.8	4.0	3.5	—	—

All plasmids used to assay promoter and terminator segments were derivatives of pXY300 (Fig. 1b). Plasmids pXY301–pXY307 contain segments from the *Hin*dIII site in gene *Q* to a *Bal*31 nuclease-generated end at the designated nucleotide near the late RNA start site (see Fig. 1a), followed by the *Eco*RI linker GGAATTCC; the number in parentheses indicates the last nucleotide that is present in the deletion. Other plasmids are described in Materials and methods. Experiments 1 and 2 are two typical sets of data giving the levels of galactokinase activity in HB101[F′*lacI*^{Q}] carrying each plasmid, and also carrying pUV5-Q2. For induced samples, IPTG was added to 1 mM to cells at an OD_{650} of 0.2 growing in LB, and growth was continued for 1.5 hr before sampling; the final OD_{650} of the cultures was 0.65–0.8. Activity is given in the units defined by McKenney et al. (1981). The last column is the result of in vitro transcription of fragments from the same plasmids, and is the data shown in Fig. 3, analyzed by excising gel slices and counting in a scintillation counter.

kinase is not inducible by IPTG when the test plasmid contains promoters from which transcription should not be affected by Q: $\lambda p_{R'}$, the major λ early rightward promoter (pXY314), and the late promoter of phage 82 (pXY312), which does not respond to λ Q function in vivo (Schechtman et al. 1980) or to purified λ Q protein in vitro (see below). Also as expected, deletion of the −35 and −10 promoter elements of $p_{R'}$ (pXY301) prevents galactokinase induction. Furthermore, galactokinase is inducible if the entire 6S RNA coding segment containing $p_{R'}$ and $t_{R'}$ is present, either alone (pXY310) or combined with t_O so that two terminators must be passed (pXY308).

About 10 times as much galactokinase (2.5–5 units) is made from plasmids containing $p_{R'}$ and its derivatives in the absence of Q protein (no IPTG added) than from a plasmid lacking a promoter altogether (pXY301). This synthesis is not caused by residual expression of Q from the repressed plasmid, because it is not dependent upon the presence of the Q source pUV5-Q2 (data not shown); it must therefore result from leakage of transcription past the terminator. This leakage allows us to verify and measure the activity of promoters in the pXY300 series plasmids in the absence of Q protein.

We cannot calculate the efficiency of antitermination in vivo directly, because we do not know what amount of galactokinase corresponds to all transcripts initiated at $p_{R'}$ reaching the galactokinase gene. However, McKenney et al. found that 2% of transcripts leak past t_O in vivo without antitermination (McKenney et al. 1981). If t_O is 98% efficient in our plasmids, then $p_{R'}$ would give 125–250 units of galactokinase in the absence of termination. Extrapolating the data of Table 1 to growth for many generations suggests that Q protein provides as much as 50–100% readthrough of t_O in vivo. We find up to 50% readthrough of t_O in the presence of purified Q protein in vitro (Table 1).

To determine how galactokinase expression from $p_{R'}$ depends on the intracellular concentration of Q protein, we grew a cell carrying the Q-source plasmid pUV5-Q2 and the Q-responsive test plasmid pXY306 in varying concentrations of IPTG, and measured Q protein by immunoblot and galactokinase by assay in cell extracts (Fig. 2a). The measurements were made many generations after IPTG was added, so that an equilibrium concentration of both Q protein and galactokinase should have been obtained. As expected, expression of galactokinase did cause the copy number of pXY306 to decrease at equilibrium, although no more than twofold at the highest level of induction (data not shown). Thus, the relative rate of galactokinase expression may be underestimated by about twofold at the highest IPTG concen-

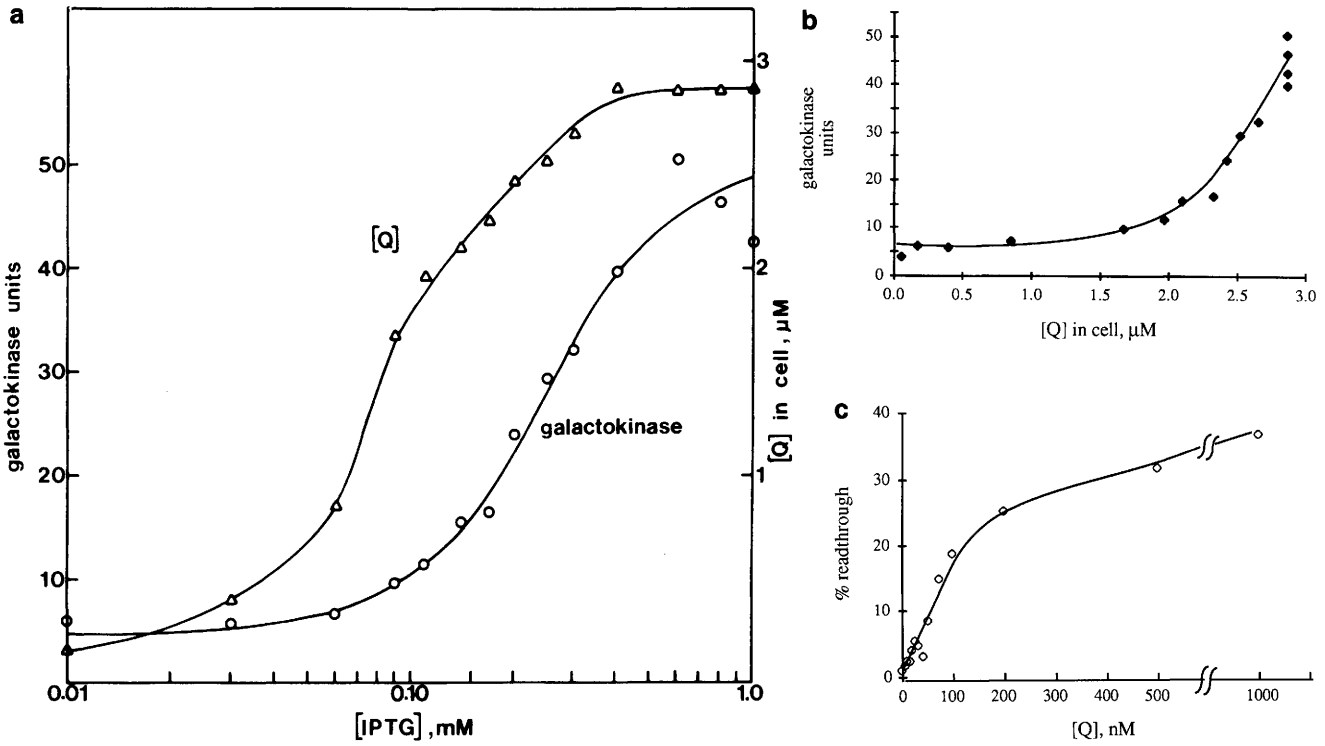

Figure 2. Concentration dependence of Q protein activity in vivo and in vitro. (*a*) Content of Q protein and activity of Q protein, in cells carrying plasmids pUV5-Q2 and pXY306, as a function of inducer concentration. *E. coli* strain HB101[F′*lacI*^{Q}] carrying pUV5-Q2 and pXY306 was grown in varying concentrations of IPTG as indicated, and samples were subjected to immunoblot assay for Q protein and to galactokinase assay. (*b*) The data above replotted to show Q activity as a function of Q protein concentration in the cell. (c) Concentration dependence of Q activity in vitro. The *Hin*dIII-*Eco*RI fragment of λ containing the natural $p_{R'}$–$t_{R'}$ segment was transcribed in the presence of increasing concentrations of purified Q protein and the percent readthrough of terminator $t_{R'}$ was determined as described (Grayhack et al. 1985).

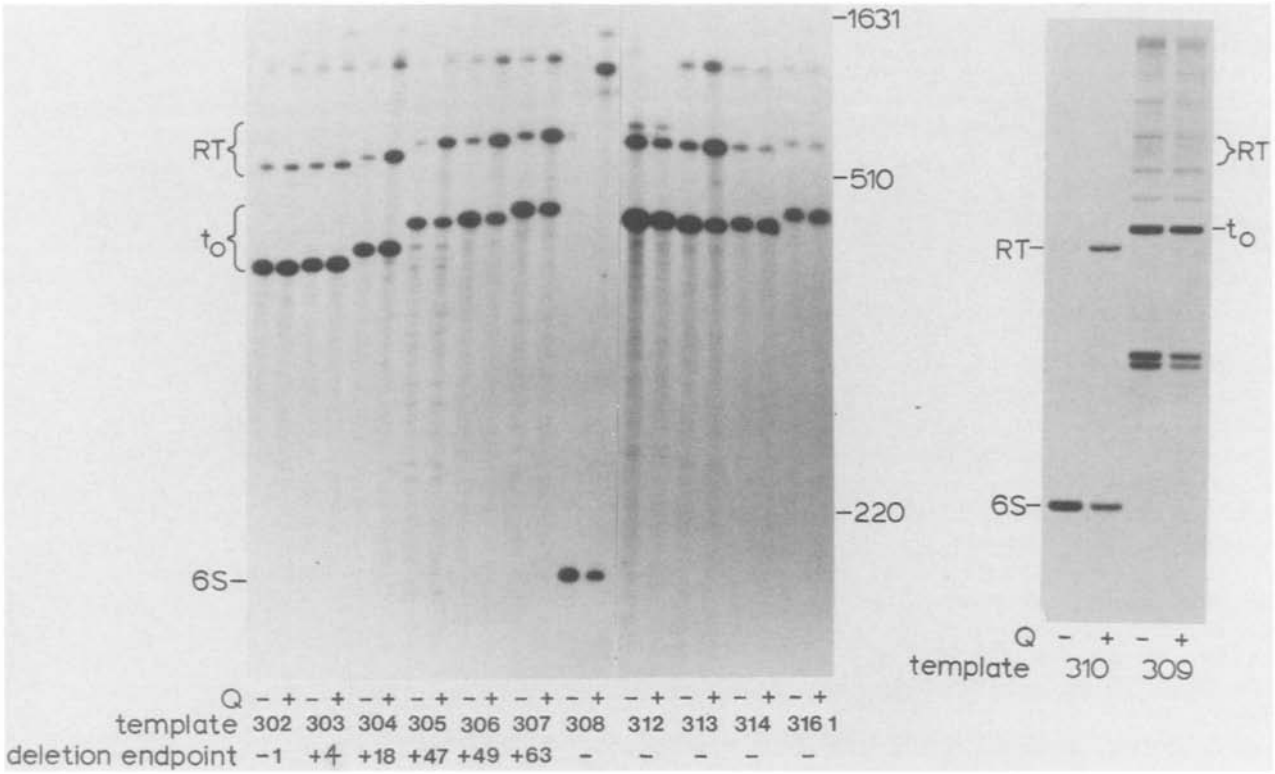

Figure 3. Activity of promoter segments in supporting antitermination by purified Q protein in vitro. Restriction fragments were transcribed and the RNA products analyzed as described in Materials and methods. RT, 6S, and t_O designate, respectively, DNAs that read through to the end of the fragment, stop at the terminator $t_{R'}$, or stop at terminator t_O. All templates except those from plasmids pXY310 and pXY309 were *Hin*dIII–*Sma*I fragments (Fig. 1). The *Hin*dIII–*Eco*RI fragment from pXY 310 is the same as the wild-type 6S RNA segment shown in Fig. 1a. The template from pXY309 was a *Hin*cII–*Sma*I fragment that also contains the bla promoter from pBR322, which gave rise to the two lower bands in the 309 lane. In this case, because of the complexity of the product in the high-molecular-weight region of the gel, it is not clear which of the larger bands is readthrough RNA from p_O, but it is clear that Q protein gives no increase in any of them.

tration, but this should not affect the shape of the curve significantly. Other experiments, in which more Q protein was made from a stronger promoter, suggest that the highest concentration of Q protein achieved in the experiment of Figure 2a is sufficient to give maximal antitermination.

We draw two conclusions from Figure 2. First, the response of galactokinase synthesis to Q protein is distinctly nonlinear (Fig. 2b). This explains the 20-min lag before galactokinase begins to accumulate after induction by IPTG. Second, more Q protein is required in vivo than in vitro. Half-maximal galactokinase synthesis in vivo requires at least 2.5 μM Q protein, whereas half-maximal antitermination in the corresponding in vitro experiment (Fig. 2c) requires only 100 nM Q protein; we discuss this disparity below.

Measurement of promoter response to Q protein in vitro

The natural target of Q protein, the segment of DNA that contains $p_{R'}$ and $t_{R'}$ and encodes the λ6S RNA (Fig.1), responds to purified Q protein in vitro (Grayhack and Roberts 1982; Grayhack et al. 1985). Efficient antitermination at $t_{R'}$ requires both Q protein and NusA protein, an *E. coli* transcription factor known also to be required for the antitermination activity of λ N protein (Friedman and Baron 1974; Greenblatt et al. 1980; Nakamura et al. 1986). We find that Q protein also is active on transcription in vitro from the pXY300 series test plasmids that combine a segment containing $p_{R'}$ with the terminator t_O. To determine the effect of purified Q protein in vitro, we take from the pXY300 series plasmid the *Hin*dIII–*Sma*I fragment containing the promoter and terminator (Fig. 1), and measure transcription reading through the terminator to the end of the fragment; Q protein increases the production of readthrough transcript when a Q-responsive promoter segment is present (e.g., pXY307). Figure 3 shows the effect of Q protein in vitro on some of the promoter segments assayed in vivo in Table 1. The smallest transcript, between 368 and 435 nucleotides in length for all of the pXY300 series except pXY308, results from termination at t_O, and the next largest transcript (508–575 nucleotides) is the readthrough to the fragment end at the *Sma*I site. (The fainter, larger transcript has not been identified, but it might be a full-length transcript of the fragment that initiates at one end.)

About 5–7% of the RNA polymerase that initiates at $\lambda p_{R'}$ in vitro reads through t_O in the absence of Q protein (but in the presence of NusA protein). The effect of Q protein is to increase readthrough to 40–55% for a promoter segment that responds to Q, for example, the fragment from pXY306 containing $p_{R'}$ and 49 nucleotides of the 6S RNA coding sequence. As we observed previously (Grayhack and Roberts 1982), response to λ Q protein in vitro is specific to $\lambda p_{R'}$; thus λ Q protein gives no increase in readthrough from DNA segments containing the λ early rightward promoter p_R or the late promoter of phage 82 (pXY314, pXY312). Figure 3 and Table 1 show that the response of each fragment to Q protein in vitro matches the response of the parental plasmid to Q protein in vivo.

As is true for the natural segment $p_{R'}$–$t_{R'}$ on which Q acts, antitermination from the reconstructed $p_{R'}$–t_O segments requires NusA protein. Figure 4 shows that NusA protein has distinct effects on transcription in vitro from this DNA in the absence and in the presence of Q protein. Without Q, NusA protein depresses readthrough of t_O from 25% to about 4%. With Q, NusA *increases* readthrough to about 40%. The effect of NusA on the natural $p_{R'}$–$t_{R'}$ segment is similar, although there is less readthrough in the absence of both Q and NusA (Grayhack et al. 1985). A second difference is that whereas Q slightly increases readthrough from $p_{R'}$–$t_{R'}$ in the absence of NusA (Grayhack et al. 1985), we have observed no effect of Q without NusA at all with $p_{R'}$–t_O, although a few percent increase in readthrough might be undetectable.

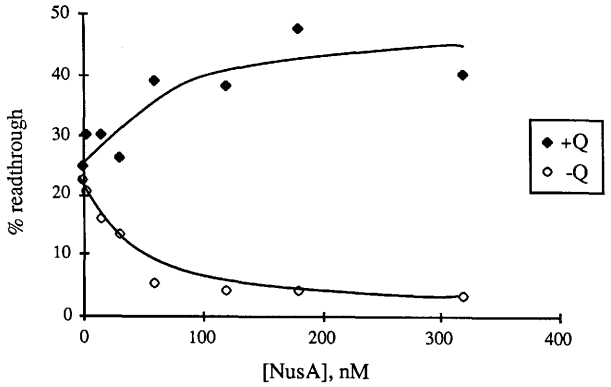

Figure 4. The effect of NusA protein on transcription termination and Q-mediated antitermination at t_O in vitro. The DNA template was the *Hin*dIII–*Sma*I fragment of pXY307. Transcription reactions contained 800 nM Q protein where indicated.

The site essential for Q function in vivo and in vitro includes the DNA segment spanning the RNA start site

To identify sequences associated with the late gene promoter that are essential for Q function, we assayed seven deletion DNAs for response to Q protein in vivo and in vitro (Fig. 3; Table 1). These deletions have removed segments of DNA from the right of the RNA start site (Fig. 1), and the remainder of the $p_{R'}$ promoter region is fused through an *Eco*RI linker to a DNA segment containing terminator t_O, located 370 nucleotides downstream of the joint. Deletions to +63, +49, +47, and +18 all respond to Q, whereas deletions to +4, −1, and −36 do not. Deletion Δ[+18] reproducibly responds to Q in vitro about half as well as the previous three. As expected, the deletion to −36 destroys promoter function in vitro (data not shown) and in vivo (Table 1), because it removes the essential promoter elements at −35 and −10. However, the deletions to −1 and +4 leave an active promoter that does not respond to Q. These results show that essential sequences exist between +4 and +18, a segment of the 6S RNA coding region. Somasekar and Szybalski (1983) also found that sequences after the RNA start site are necessary for Q function in vivo.

To test whether the segment from +4 to +18 is sufficient for Q function, we fused it to a different promoter. Using the *Hin*fI site GACTC beginning at nucleotide −5 of $p_{R'}$ and a *Hin*fI site beginning at −7 of the phage λ promoter p_O, we constructed the fusion pXY309 shown in Figure 1; it contains the $p_{R'}$ segment from −6 to +47 and λp_O sequences to the left of −6. This is an active promoter in vivo and in vitro (Fig. 3, Table 1), but it does not respond to Q by either assay. The fusion is not perfect, in that the $p_{R'}$ segment is one nucleotide closer to the −10 promoter element than it is in $p_{R'}$ itself; it is possible that the RNA initiates at the second adenine or at another nucleotide nearby. Despite this reservation, we conclude that the presence of the −6 to +47 segment in nearly its natural location does not confer on a foreign promoter the ability to respond to Q protein.

Two further modifications of the $p_{R'}$ segment show that the mere presence of all of the natural sequences before and after the RNA start site does not suffice to confer response to Q protein, if these segments are separated. We made insertions of 10 and 18 nucleotides between −1 and +1 of $p_{R'}$, yielding plasmids pXY316-1 and pXY316-2 (Fig. 1). By the previous criteria, both are active promoters, but neither responds to Q protein in vivo or in vitro (Fig. 3, Table 1). The behavior both of these altered promoters and of the deletions indicates that a DNA segment including the promoter, and a portion of the first 18 nucleotides of the 6S coding sequence must be intact for Q to modify transcription.

We showed previously that the transcription pause at +16 occurs normally during in vitro transcription of DNA of the Q-responsive plasmid pXY304 [Δ(+18)], but not DNA of the nonresponsive plasmid pXY302 Δ[(−1)] (Grayhack et al. 1985). We also have found that the other derivatives of $p_{R'}$ that do not respond to Q—the deletion pXY303 [Δ(+4)], the insertion mutants pXY316-1 and pXY316-2, and the $p_O/p_{R'}$ fusion pXY309—also do not encode a detectable pause at +16 during transcription in vitro (X. Yang and J.W. Roberts, unpubl.). Furthermore, the insertion mutants do not encode pauses displaced downstream by the length of the insertion, as might be expected if the pause were encoded only in the originally transcribed DNA segment that is moved downstream by the insertion. Our results imply that pausing at +16 is

affected by sequences both upstream of −6 and downstream of +4, and that these sequences cannot be separated by 10 or 18 bp of extraneous DNA. The failure of all the mutant DNAs to encode a pause at +16 may be sufficient to explain their inability to respond to Q protein.

High efficiency of antitermination at the second of two tandem terminators

The greatest efficiency of antitermination that we have found in vitro is about 50%. By constructing plasmids containing two terminators in a row and measuring readthrough of the second terminator, we can ask if this efficiency reflects the fraction of RNA polymerase molecules that are modified by Q, or, instead, the probability that a Q protein-modified RNA polymerase escapes a terminator. If antitermination by the modified RNA polymerase is 50% efficient at each terminator, then only about half of the transcripts that pass the first terminator should also pass the second. On the other hand, if antitermination is more nearly 100% efficient, but only half of the RNA polymerase molecules are modified by Q near the promoter, then most of the enzyme molecules that pass the first terminator also should pass the second.

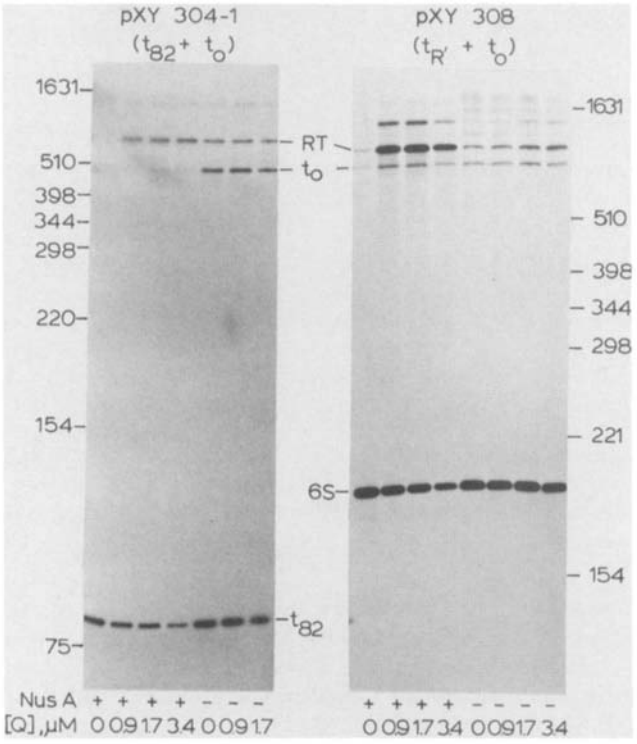

Figure 5. Activity of Q protein on tandem terminators in vitro. Templates were *Hin*dIII–*Sma*I fragments of the indicated plasmids. Plasmid pXY304-1 is derived from pXY304 (the +18 deletion) by insertion, into its *Eco*RI site, of a 92-bp fragment containing the terminator of phage 82, which corresponds to $t_{R'}$ of λ (J. Goliger and J.W. Roberts, unpublished). RNA initiated at $p_{R'}$ and terminated at the phage 82 terminator is designated t_{82} in the figure; RNA that reads through the phage 82 terminator to the second terminator is designated t_O in the figure. RNA reading through both terminators is designated RT. Plasmid pXY308 has the wild-type 6S RNA region (the *Hin*dIII–*Eco*RI segment of Fig. 1a) cloned into the pXY300 series vector described in Fig. 1b, so that promoter $p_{R'}$ is followed by terminators $t_{R'}$ and t_O in tandem. RNA initiated at $p_{R'}$ and terminated at $t_{R'}$ is named 6S; RNA that reads through $t_{R'}$ to the second terminator is designated t_O. NusA protein was added to 150 nm where present.

The behavior of plasmids containing two terminators in a row, pXY304-1 and pXY308, shows that the second possibility is correct. Plasmid pXY304-1 is identical to pXY304, except that a DNA segment containing the terminator t_{82} (the counterpart of $\lambda t_{R'}$, derived from phage 82) is placed in the *Eco*RI site at +18 between $p_{R'}$ and terminator t_O. Plasmid pXY308 has the natural $p_{R'}$–$t_{R'}$ segment placed before terminator t_O. Figure 5 shows the effect of Q protein and NusA protein on in vitro transcription of these plasmids; each RNA band is designated by the name of the terminator that gives rise to it. In the presence of NusA alone, little synthesis escapes the first terminator in either plasmid. With NusA and Q protein together, most RNA that escapes the first terminator reaches the end of the fragment and does not stop at t_O, the second terminator. The effect is particularly apparent for pXY304-1, because in this experiment the Δ(+18) deletion allows only about 15% antitermination through the first terminator, as we determined by counting radioactivity in pieces of the gel of Figure 5 directly (see also Table 1); yet 12% of the RNA polymerase goes through both terminators. The result of omitting NusA protein confirms that there is little antitermination by Q protein without NusA. It is useful that enough transcription leaks past t_{82} in the absence of NusA to mark the position of the transcript terminated at t_O in pXY304-1, and thus to demonstrate that t_O is in fact functional in this plasmid.

The activity of pXY308 in vivo is consistent with these results (Table 1). Both terminators are active in vivo, because galactokinase synthesis without IPTG is 10-fold less than with either terminator alone (compare pXY308 with pXY307 and pXY310), and is in fact equal to the promoter-less background (pXY301). Yet induction of Q protein with IPTG yields only slightly less galactokinase from pXY308 than from plasmids with only one terminator.

Discussion

These experiments define the site at which the phage λ gene Q antiterminator modifies RNA polymerase. The smallest DNA segment we isolated that responds to Q protein includes the −35 and −10 elements of the late gene promoter and extends through 18 nucleotides of the late RNA coding region. The inability of several rearrangements of this segment that are still active promoters to respond to Q suggests that Q function requires the natural sequence spanning the late RNA start, extending from a site upstream of nucleotide −6 to a site between nucleotides +4 and +18. Since Q protein can modify RNA polymerase that has paused after making

16 nucleotides of the late gene transcript, pausing at +16 may be a necessary step in the engagement of Q protein. Correspondingly, the DNA segment essential for Q function may include sequences that induce the pause at +16, as well as sequences that Q protein binds. All of the Q-nonresponsive promoters we have made by deletion or rearrangement of $p_{R'}$ fail to encode the pause, and this defect may be sufficient to account for their inactivity. Thus, these experiments do not show which DNA segments in the essential region Q protein binds (or which DNA segments encode RNA that Q protein binds). We speculate that when RNA polymerase contacts DNA both in the promoter and in the late RNA coding region, a pause ensues; and this pause provides the substrate that Q protein recognizes.

The complete agreement between in vivo and in vitro assays of each DNA for response to Q protein gives us confidence that the in vitro reaction accurately reflects the action of Q protein in vivo. Reconstructing the regulatory system on plasmids also allowed us to show that two terminators foreign to the natural λ late gene regulatory segment, t_{O} and t_{82}, are passed by Q-modified RNA polymerase. This result agrees with the notion that the genome specificity of antitermination resides in the Q protein engagement site near the promoter, and not in the terminator.

We designed the Q-source plasmid (UV5-Q2) through trial-and-error to regulate $p_{R'}$ properly. A similar plasmid containing only one copy of the *Q* gene transcribed from promoter *lacUV5* gave a very low induced level of galactokinase, whereas another plasmid carrying one copy of *Q* transcribed from the stronger *tac* promoter leaked enough Q activity to give a high level of galactokinase from responsive plasmids without induction (X. Yang and J.W. Roberts, unpubl.). The relative level of expression of the *Q* gene from the three plasmids is consistent with the measured strength of the *lacUV5* and *tac* promoters.

Why might more Q protein be required in vivo than in vitro? Although the Q protein concentration in the cell is somewhat uncertain, the determination is unlikely to be wrong by a factor of 25. The protein itself should not be different, because the DNA segment carrying gene *Q* in UV5-Q2 is also the source of Q protein in a different plasmid used to make Q protein for purification. Possibly most Q protein in the cell is unavailable because it is insoluble or trapped by incorrect binding. Conceivably Q protein exists in a form more favorable to its function after it has been subjected to purification, e.g., as a dimer that assembles much less efficiently in vivo. A more interesting explanation is that a high concentration of Q protein is required for kinetic reasons. The transcriptional pause at +16 where Q acts is very long in vitro, lasting about 5 min. Presumably Q protein cannot modify an RNA polymerase molecule that has escaped the pause (Grayhack et al. 1985); if the pause were much shorter in vivo [and it is known that some pauses are unnaturally long in vitro (Kassavetis and Chamberlin, 1981)], then a much higher concentration of Q protein could be required for it to react efficiently with RNA polymerase during the pause in the cell. During phage growth, a high concentration might be unnecessary if Q protein is made near its site of action and thus can act in *cis*. Q is thought to act more efficiently in *cis* than in *trans* in vivo (Echols et al. 1976; Burt and Brammar 1982).

The fact that antitermination is only partly efficient in vitro at the first of two terminators, but almost completely efficient at the second, also may be explained by the behavior of Q protein at the +16 pause site. In vitro, only about half of the RNA polymerase molecules pause at all; presumably the 50% antitermination efficiency reflects the fraction that do pause, and are therefore susceptible to modification by Q protein. Once it is modified, the ability of RNA polymerase to pass one or several terminators is then very high.

Materials and methods

Bacterial strains

All experiments used the strain HB101[F'*lacI*^{Q}]; *galK2*, *recA13*, obtained from K. McKenney.

Plasmids

pUV5-Q2. Plasmid pUV5-Q2 was derived from the low-copy-number plasmid pSE150, which is compatible with pBR322 and which confers spectinomycin resistance (S.J. Elledge and G.C. Walker, pers. comm.). First, a *Bam*HI–*Eco*RI fragment containing the *lacUV5* promoter fused to gene *Q* [deleted to the right end of the gene at nucleotide 44,549 (Daniels et al. 1983) and ended with an *Eco*RI linker] was cloned into the 5.5-kb *Eco*RI–*Bam*HI fragment of pSE150; then a second copy of the *Q* gene, bounded on the right by an *Eco*RI site as above, and on the left by the *Cla*I site at 43,825 filled and ended with an *Eco*RI linker, was inserted into the *Eco*RI site.

pXY300 series. These were derived from pKO1 (McKenney et al. 1981). The unique *Hin*dIII site of pKO1 was changed to a *Bgl*II site, and a 190-bp *Sau*3A1 fragment containing the λt_{O} terminator (bp 38,476–38,664; Daniels et al. 1983) was cloned into the *Bgl*II site. A plasmid with λt_{O} in the correct orientation was chosen. The unique *Eco*RI site 300 bp upstream of the t_{O} insertion site was opened, an *Eco*RI–*Bam*HI adapter added to the left side (as oriented in Fig. 1), and the 375-bp *Bam*H1–*Hin*dIII piece from pBR322 added to the *Bam*H1 adapter end; the *Eco*RI site to the left of *Bam*HI site was then destroyed. All of the pXY300 series plasmids were constructed by inserting either a *Hin*dIII–*Eco*RI fragment or a *Bgl*II–*Eco*RI fragment into this vector.

Plasmids pXY301–pXY307 contain segments from the *Hin*dIII site in gene *Q* to a *Bal*31 nuclease generated end at the designated nucleotide near the late RNA start site (see Fig. 1a), followed by the *Eco*RI linker GGAATTCC; the number in parentheses indicates the last nucleotide that is present in the deletion. Plasmid pXY310 contains the wild-type *Hin*dIII–*Eco*RI piece (Fig. 1a) cloned into a derivative of pXY300 from which terminator t_{O} had been deleted. Plasmid pXY308 is described in the legend to Figure 5. Plasmid pXY313 contains the late RNA promoter region of phage P22; an *Ava*I site at the +48 position of the P22 late RNA coding region was changed to an *Eco*RI site, and a 480-bp fragment containing the carboxyterminal part of P22 Q protein (bounded by a *Hin*dIII site on the left), the P22 late promoter, and the first 48 bases of the P22 late

RNA was cloned into the pXY300 vector. Plasmid pXY312 contains the late promoter and the first 52 bases of the late RNA of phage 82 in a 160-bp *Hin*dIII–*Eco*RI fragment (J. Goliger and J.W. Roberts, unpubl.). pXY314 contains the λ early promoter p_R and the first 50 bases of the Cro message as described (Lau and Roberts 1985).

Variant pXY309 was constructed as follows: the 77-bp *Bgl*II–*Hin*fI DNA fragment of λ (nucleotides 38,678–38,755; Daniels et al. 1983) containing the −35 and −10 sequences of promoter p_O was ligated to the 57-bp *Hin*fI–*Eco*RI fragment from pXY305 (containing nucleotides −4 to +47 of the wild type p_R sequence, plus GGAATT from the *Eco*RI linker), and the resulting *Bgl*II–*Eco*RI fragment was cloned into the *Bam*HI–*Eco*RI sector of the pXY300 series plasmids. Variants pXY316-1 and pXY316-2 were constructed as follows: the *Hin*dIII–*Eco*RI piece of pXY302 (the deletion to −1), blunt-ended by filling its *Eco*RI end, was ligated in the presence of *Pst*I linker (GCTGCAGC) to the *Hin*fI–*Eco*RI piece of pXY306 (nucleotides −4 to +49 of the wild-type sequence, plus GGAATT from the *Eco*RI linker), which had been filled at its *Hin*fI end. The resulting *Hin*dIII–*Eco*RI pieces were cloned into the *Hin*dIII–*Eco*RI sector of the pXY300 series plasmids. The resulting plasmid pXY316-1 received one *Pst*I linker during the blunt-end ligation and thus has a 18-bp insert containing a *Pst*I site, whereas pXY316-2 received no linker and thus has a 10-bp insert.

In vitro transcription

Transcription in vitro was performed and transcripts were analyzed by gel electrophoresis as described (Grayhack et al. 1985). Synthesis was initiated by simultaneous addition of magnesium ions and rifampicin. All reaction mixtures contained 150 nM NusA protein and 500 nM Q protein (where present), unless indicated otherwise. DNA fragments were purified by precipitation of larger fragments with polyethylene glycol (Lis 1980).

Measurement of Q protein mass in cells

E. coli strain HB101[F′*lacI*^{Q}] carrying pUV5-Q2 and pXY306 was grown in L broth containing 0.04 mg/ml spectinomycin, 0.1 mg/ml ampicillin, and different concentrations of IPTG. After growth to OD_{650} = 0.9, samples were subjected to immunoblot assay for Q protein as described by Lovett and Roberts (1985). The amount of Q protein in aliquots of cells was determined by scanning with a densitometer SDS gel transfers stained as described by Lovett and Roberts (1985), using as a standard purified Q protein mixed with cells and processed through electrophoresis and blotting exactly like the experimental samples; the concentration of Q protein was calculated assuming a cell volume of 1 μ^3.

Assay of galactokinase

Galactokinase was assayed as described (McKenney et al. 1983).

Acknowledgments

These experiments were supported by grant GM21941 from the National Institutes of Health. We thank John Lis for criticizing the manuscript, M.J. Morse for constructing plasmid pXY309, K. McKenney, G. Walker, and S. Elledge for strains, and J. Goliger for helpful discussions and gifts of materials.

References

Bentley, D.L. and M. Groudine. 1986. A block to elongation is largely responsible for decreased transcription of c-*myc* in differentiated HL60 cells. *Nature* **321**: 702–706.

Burt, E. and W. Brammar. 1982. The *cis* specificity of the Q gene product of bacteriophage lambda. *Mol. Gen. Genet.* **185:** 486–472.

Daniels, D., J. Schroeder, W. Szybalski, F. Sanger, A. Coulson, G. Hong, D. Hill, G. Petersen, and F. Blattner. 1983. Complete annotated lambda sequence. In *Lambda II* (ed. R. Hendrix, J. Roberts, F. Stahl, and R. Weisberg), pp. 519-676. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

deCrombrugghe, B., M. Mudryj, R. Di Lauro, and M. Gottesman. 1979. Specificity of the bacteriophage lambda *N* gene product (N): *Nut* sequences are necessary and sufficient for the antitermination by N. *Cell* **18:** 1145–1151.

Echols, H., D. Court, and L. Green. 1976. On the nature of *cis*-acting regulatory proteins and genetic organization in bacteriophage: The example of gene Q of bacteriophage lambda. *Genetics* **83:** 5–10.

Friedman, D. and L. Baron. 1974. Genetic characterization of a bacterial locus involved in the activity of the *N* function of phage λ. *Virology* **58:** 141–148.

Friedman, D. and M. Gottesman. 1983. Lytic mode of lambda development. In *Lambda II* (ed. R. Hendrix, J. Roberts, F. Stahl, and R. Weisberg), pp. 21-51. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Gilmour, D. and J. Lis. 1986. RNA polymerase II interacts with the promoter region of the noninduced *hsp70* gene in *Drosophila melanogaster* cells. *Mol. Cell. Biol.* **6:** 3984–3989.

Grayhack, E. and J. Roberts. 1982. The phage λ gene Q product: Activity of a transcription antiterminator in vitro. *Cell* **30:** 637–648.

Grayhack, E.J., X. Yang, L.F. Lau, and J.W. Roberts. 1985. Phage lambda gene *Q* antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site. *Cell* **42:** 259–269.

Greenblatt, J., J. Li, S. Adhya, D.I. Friedman, L.S. Baron, B. Redfield, H. Kung, and H. Weissbach. 1980. L factor that is required for β-galactosidase synthesis is the *nusA* gene product involved in transcription termination. *Proc. Natl. Acad. Sci.* **77:** 1991–1994.

Kassavetis, G. and M. Chamberlin. 1981. Pausing and termination of transcription within the early region of bacteriophage T7 DNA in vitro. *J. Biol. Chem.* **256:** 2777–2786.

Lau, L.F. and J. Roberts. 1985. ρ-Dependent transcription termination at λ_{R1} requires upstream sequences. *J. Biol. Chem.* **260:** 574–584.

Lis, J. 1980. Fractionation of DNA fragments by polyethylene glycol induced precipitation. *Methods Enzymol.* **65:** 347–353.

Lovett, C. and J. Roberts. 1985. Purification of a RecA protein analogue from *Bacillus subtilis*. *J. Biol. Chem.* **260:** 3305–3313.

McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady, and M. Rosenberg. 1981. A system to study promoter and terminator signals recognized by *Escherichia coli* RNA polymerase. In *Gene amplification and analysis*, Vol. II. *Analysis of nucleic acids by enzymatic methods* (ed. J.G. Chirikjian and T.S. Papas). Elsevier-North Holland, Amsterdam.

Nakamura, Y., S. Mizusawa, D. Court, and A. Tsugawa. 1986. Regulatory defects of a conditionally lethal *nusA*ts mutant of *Escherichia coli*: Positive and negative modulator roles of NusA protein *in vivo*. *J. Mol. Biol.* **189:** 103–111.

Salstrom, J.S. and W. Szybalski. 1978. Coliphage λ *nut*L$^-$: A unique class of mutants defective in the site of gene N product utilization for antitermination of leftward transcription. *J. Mol. Biol.* **134:** 195–221.

Schechtman, M.G., J. Snedeker, and J.W. Roberts. 1980. Genetics and structure of the late gene regulatory region of phage 82. *Virology* **105:** 393–404.

Somasekhar, G. and W. Szybalski. 1983. Mapping of the Q-utilization site (*qut*) required for antitermination of late transcription in bacteriophage lambda. *Gene* **26:** 291–294.

Transcription antitermination by phage lambda gene Q protein requires a DNA segment spanning the RNA start site.

X J Yang, C M Hart, E J Grayhack, et al.

Genes Dev. 1987, **1:**
Access the most recent version at doi:10.1101/gad.1.3.217