Studying Promoters and Terminators by Gene Fusion

Martin Rosenberg, Ana B. Chepelinsky, Keith McKenney

Summary. Prokaryotic gene control signals can be isolated, compared, and characterized by precise fusion in vitro to the *Escherichia coli* galactokinase gene (*galK*), which provides both a simple assay and genetic selection. This recombinant *galK* fusion vector system was applied to the study of promoters and terminators recognized by the *Escherichia coli* RNA polymerase. Three promoters created by mutation from DNA sequences having no promoter function were characterized. Mutations that inactivate promoter function were selected, structurally defined, and functionally analyzed. Similarly, transcription termination was examined, and mutations affecting terminator function were isolated and characterized.

Gene fusion techniques have proved extremely useful in the study of prokaryotic gene regulation (*1–10*). These techniques have depended on the genetic fusion of an assayable, selectable gene function into the operon of interest. The fusion places the gene under the transcriptional control of that operon, thereby allowing study of the operon's regulatory information. Although genetic fusions have expanded our knowledge of operon control elements, these techniques have limitations. More recently, the concept of gene fusion has been combined with recombinant technology for the study of transcriptional regulatory elements (*11–15*). DNA fragments containing specific transcription control regions are fused in vitro to a gene function that can be readily assayed and genetically selected. This is best done in a recombinant vector system designed and constructed to circumvent the limitations inherent in the conventional genetic fusion systems so that gene regulatory elements can be studied in new ways.

M. Rosenberg is on the staff of the Laboratory of Biochemistry, National Cancer Institute, Bethesda, Maryland 20205, and is currently the director of the Department of Molecular Genetics, Smith Kline & French Laboratories, Philadelphia, Pennsylvania 19101. At the time of this study, A. B. Chepelinsky and K. McKenney were working in the Laboratory of Biochemistry, National Cancer Institute. A. B. Chepelinsky is now in the Department of Molecular and Developmental Biology, the National Eye Institute, Bethesda 20205, and K. McKenney is at the Laboratory of Molecular Biology, Medical Research Council Centre, University Medical School, Cambridge CB2 2QH, England. Correspondence should be addressed to M. Rosenberg.

We earlier described a recombinant vector system developed to isolate, compare, and characterize almost any promoter or terminator signal recognized by the *Escherichia coli* RNA polymerase (*14, 15*). In this system the *E. coli* galactokinase gene (*galK*) provides the assayable, selectable function to which transcription regulatory signals are fused. There is a simple, sensitive assay for galactokinase and its expression can be made either essential or lethal to cells under the appropriate growth conditions (that is, *galK* can be selected either for or against). This dual selection system for studying gene regulatory elements utilizes plasmid, phage, and bacterial vectors, and it provides flexibility in that any construction made in the multicopy plasmid vector system can be moved to the phage vector or selectively into the bacterial chromosome. This allows regulatory signals to be studied either in single- or multicopy and in either the episomal or chromosomal state. The system is being used to study various prokaryotic regulatory signals (*16–26*). We recently expanded the application of the *galK* fusion system to study eukaryotic gene control elements in yeast (*27, 28*) and in higher cell systems (*15, 29–31*).

In this article, we describe several applications of our *galK* fusion vector system to the study of prokaryotic promoters and terminators. We show how the system can be used to characterize promoter signals created by mutation, to select and functionally characterize mutations that inactivate promotion, and to isolate and characterize mutations in transcription termination signals. These studies emphasize the utility of the fusion vector system for quantitative assessment of regulatory signal function both in vivo and in vitro. Other aspects of prokaryotic transcriptional regulatory signal structure and function have been reviewed (*32–35*).

Promoters Created by Mutation

The plasmid vector system, pKO, was used for isolating and studying promoters (Fig. 1A). Details of the design and construction of this vector system have been described (*14, 15, 36*). The *E. coli galK* gene is inserted into the vector in such a manner that it is not expressed (that is, not transcribed). Thus, the pKO plasmids will not complement a bacterial host cell deficient in *galK* expression (*galK*$^-$). DNA fragments are inserted into the vector at any of the restriction enzyme sites (R sites) that have been engineered into the vector upstream from *galK* (Fig. 1A). Insertion of a promoter signal in proper orientation results in *galK* expression (Fig. 1B). This expression is readily seen either by selective growth in the appropriate media or in color indicator plates (Table 1). Moreover, the precise levels of expression are monitored by *galK* enzyme assay. Various natural promoter signals have been isolated, compared, and characterized with this vector system (Table 1) (*14, 15, 36*). In this section we describe the cloning and characterization of three promoter signals created by mutation from DNA sequences that previously had no promoter function.

Promoter P_{482}. The first such promoter was isolated from a λ transducing phage, λ482, which carries a portion of the *E. coli* galactose operon fused to phage sequences (*37*). This particular fusion was generated within the bacterial chromosome by deleting sequences between the *gal* operon and an integrated copy of the phage. The deletion [25 kilobases (kb)] fused sequences of the λN gene to sequences in the middle of the *gal* transferase gene (*galT*). Upon cloning DNA fragments derived from the λ482 phage into the pKO vector, we repeatedly found a small fragment (0.5 kb) that activated *galK* expression and that was derived from the fusion region. This was unexpected since no promoters were known to be located in this region. The promoter activity of this fragment, called P_{482}, was measured in vivo and compared to a variety of other pKO

constructions carrying either natural or mutant promoter signals (Table 1). P_{482} is a relatively weak signal in vivo, functioning at only 20 percent of the efficiency of the more typical bacterial operon promoters such as the *lac* and *gal* promoters.

In order to position precisely the P_{482} signal on the 0.5-kb fragment, we carried out transcription studies in vitro using plasmid DNA as template. The template was cut at an appropriate restriction enzyme site positioned downstream from the P_{482} region and "run-off" transcripts labeled with ^{32}P were generated in standard transcription reactions (Fig. 2). The RNA products were resolved on polyacrylamide gels, eluted from the gel, and further analyzed by standard two-dimensional T_1-oligonucleotide analysis (not shown) (*38*). The results indicate that only a single transcription start occurs on the λ482 fragment, and this start site can be unambiguously positioned at the A residue shown in Fig. 3A. This start defines the P_{482} promoter and demonstrates that the promoter region spans the fusion junction between the λN sequence and the *galT* sequence. Comparison of the λ482 sequence with the corresponding regions of λ+ and the *gal* operon (Fig. 3A) indicates that the 25-kb deletion that generated P_{482} occurred by homologous recombination in an identical 8–base pair (bp) sequence shared by the λN gene and the *galT* gene. The λ482 fusion junction retains this 8-bp sequence with λN gene sequences on one side and *galT* sequences on the other. The P_{482} start site occurs within the λN gene sequence adjacent to the fusion junction.

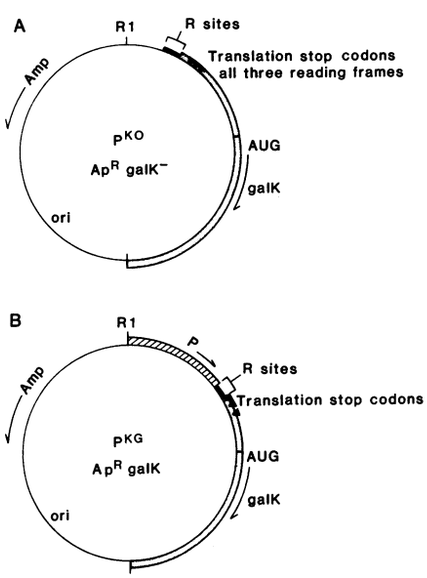

Fig. 1. (A) The pKO plasmid vector for cloning promoters. The stipled area represents the galactokinase gene (*galK*) and 168 base pairs of "leader" preceding the AUG initiation signal of *galK*. Translation stop codons were engineered into all three reading frames preventing any translation originating in the cloned insert from reaching the *galK* translation initiation signal (*14, 36*). Various restriction endonuclease sites (*R*) can be used to insert DNA fragments. There now exist at least ten pKO derivatives, which differ only in the number and types of restriction sites placed at this position. Ap^R denotes the β-lactamase gene, which provides a second selectable marker on all pKO derivatives. The replication origin (*ori*) was derived from pBR322. (B) The pKG plasmid vector for cloning terminators. This vector contains a promoter (*P*) positioned upstream from *galK*. Transcription termination sites are inserted between *P* and *galK* by means of the restriction sites (*R*) provided. All other designations are as in (A) above.

Prokaryotic promoters generally exhibit two regions of strong sequence homology, positioned 10 bp and 35 bp upstream from the transcription start site (the −10 and −35 regions, respectively) (*32–34*). Various tests have shown that the sequences in these regions and the conserved distance between them (17 bp) are important for promoter function (*32, 33*). The −10 region of P_{482} is derived from the λN gene sequence and exhibits perfect homology to the consensus sequence of the −10 region hexamer of *E. coli* promoters. In contrast, the −35 region of P_{482} is derived from the *galT* sequence and exhibits some—although rather poor—homology to the −35 hexamer consensus sequence. Apparently, the 25-kb deletion created the P_{482} promoter by fusing the −10 and −35 region sequences together at a distance appropriate to achieve promoter function. The rather poor efficiency of this promoter probably reflects the lack of a good −35 region recognition sequence for RNA polymerase.

We also examined the ability of the corresponding regions of the λ+ N gene and the *E. coli galT* gene, separately, to exhibit promoter function. Transcription experiments (Fig. 2) and RNA polymerase filter binding assays (not shown) indicated that these regions had no promoter function. In particular, the N gene sequence (Fig. 3A) is identical with P_{482} up to position −28 of the promoter region. Thus, P_{482} and N share the same consensus −10 region sequence and the same potential start-site information. It must be the −35 region sequence of P_{482}, supplied by *galT*, which selectively allows P_{482} (and not N) to function as a promoter. Comparison of the two different −35 region sequences does not distinguish why P_{482} functions and N does not. There remain subtleties inherent in the structures of these regions that RNA polymerase can discern.

P_{c17} and P_{cin}. The pKO system was used to characterize two other promoters created by mutational events. One of these, the λ*c*17 promoter (P_{c17}) arises from an exact tandem duplication of a 9-bp sequence in a virulent mutant derivative of the phage (*39*). This duplication (Fig. 3B) creates a perfect −10 region hexamer consensus sequence at the junction of the repeat. Transcription studies in vitro show that RNA polymerase recognizes this region and initiates transcription 6 bp downstream from this hexamer at the indicated C residue (Fig. 3B) (*39*).

A small DNA fragment carrying the

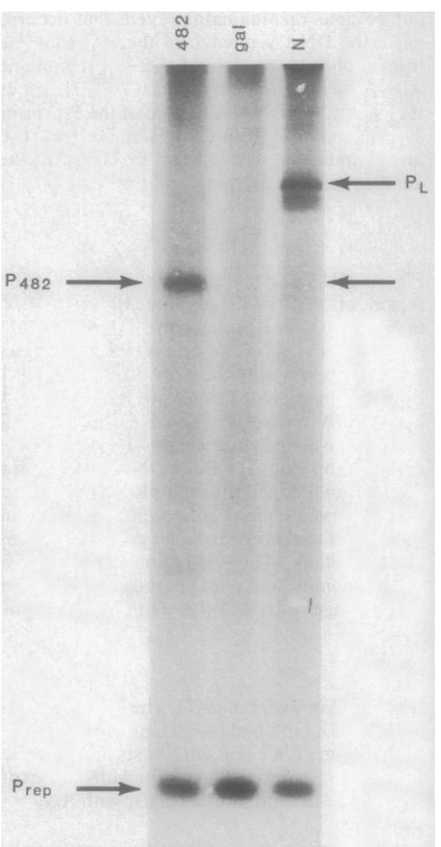

Fig. 2. Autoradiogram of a polyacrylamide gel fractionation of ^{32}P-labeled RNA synthesized in vitro from DNA templates containing the P_{482} promoter region (*482*), and the corresponding regions of the wild-type *galT* gene (*gal*) and the wild-type λN gene (*N*). Each DNA template was linearized by restriction enzyme digestion and then transcribed in vitro by standard procedures (*39*). The P_{482} run-off transcript is 250 bases in length and was identified by two-dimensional T_1-oligonucleotide fingerprint analysis (*38*). If the corresponding *galT* gene sequence of the λN gene sequence promoted transcription, then similarly sized RNA transcripts would have been obtained. No such RNA's were observed (←). The 110-base RNA from the P_{rep} promoter (*58*) serves as an internal standard and size marker. Also shown is the 320-nucleotide run-off transcript initiated from the phage λ promoter, P_L, which also occurs on the vector carrying the N gene sequence.

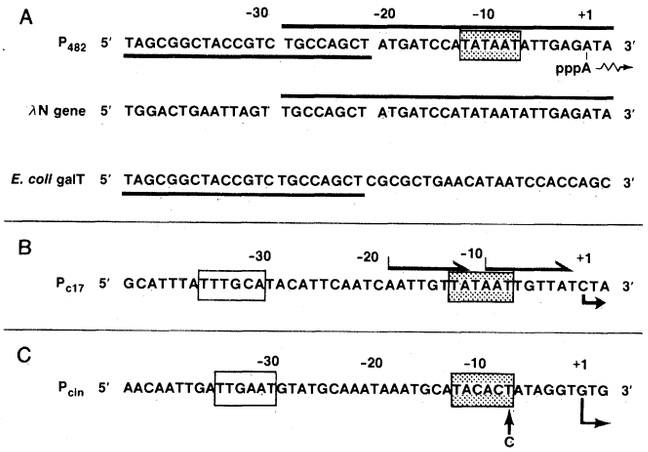

Fig. 3. Promoters created by mutation. (A) DNA sequence of the P_{482} promoter region derived from the λ482 phage (see text for details). Also shown are the DNA sequences of the corresponding regions of the N gene of phage λ (λ*N*) and the *E. coli galT* gene. The P_{482} transcription start is indicated (arrow), as is the −10 region sequence (stipled). That part of the P_{482} sequence which is in common with the N gene is overlined, and that part which is in common with *galT* is underlined. An 8-bp homology is shared by all three sequences. The P_{482} promoter was created by a 25-kb deletion resulting from a homologous recombination event that occurred between this identical 8-bp sequence in N and *galT*. (B) DNA sequence of the P_{c17} promoter region derived from the λ*c*17 phage (*39*). The transcription start site and the −10 region and −35 region consensus sequences are indicated. Also shown is the 9-bp duplication (→) that creates the −10 region sequence at the junction of the repeat. (C) DNA sequence of the P_{cin} promoter region derived from the λ*cin* phage (*39, 40*). The base change from C to T, which creates this promoter, is indicated. All other designations are as indicated above. The symbols for the nucleic acids are A, adenine; C, cytosine; G, guanine; and T, thymine.

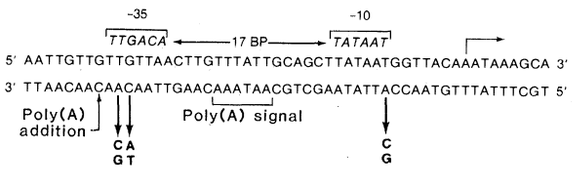

Fig. 4. DNA sequence of the region of the SV40 genome that functions as a prokaryotic promoter, P_{SV40} (*42, 44*). This region also contains the polyadenylation signal for SV40 late mRNA synthesis. The region was isolated on a single-copy 237-bp Bam HI–Bcl I DNA restriction fragment (0.15 to 0.19 map units) from the SV40 genome and cloned into the pKO vector. The P_{SV40} transcription start site is indicated, as are the conserved regions of homology with prokaryotic promoters (the −10 and −35 regions). The location of the three promoter down mutations isolated by use of the pKO system is shown (see text for details). The −35 region transition of G · C to A · T was obtained independently by both hydroxylamine and *mut*D mutagenesis. The other −35 region mutation was obtained by nitrous acid mutagenesis, and the −10 region mutation was obtained with *mut*D.

Table 1. Promoters in pKO. The construction of the vectors is described in detail in the text and in (*14*), (*15*), (*36*), and (*44*). Hosts used were *E. coli* N100 (*galK*$^-$; *recA*$^-$) and *E. coli* S165; other appropriate *galK*$^-$ and *gal*$^-$ hosts are described in (*14*) and (*44*). Media are described in detail in (*14*) and (*45*). Activity units for *galK* are expressed as nanomoles of galactose phosphorylated per minute per milliliter of cells, with A_{650} = 1 (*14*). Symbols: +, growth; −, no growth; N.D., not done.

Plasmid	*galK*$^-$ Host in		*gal*$^-$ Host in LB + galactose	*galK* Activity
	Indicator	Minimal medium		
pKO	White	−	++	10
pKO-482	Red	+	+−	120
pKM-1	White	−	++	10
pKO-*c*17	Red	+	−	350
pKM-2	White	−	++	10
pKO-*cin*	Red	+	−	700
pKO-*gal*	Red	+	−	650
pKO-*lac*	Red	+	−	500
pKO-*tet*	Red	+	−	550
pKO-SV40	Red	+	−	220
pKO-SV40M1	N.D.	N.D.	+	88
pKO-SV40M2	N.D.	N.D.	+	85
pKO-SV40M3	N.D.	N.D.	+	75

λ*c*17 promoter and the corresponding fragment from λ+ was cloned into the pKO system. The resulting vectors, pKO-*c*17 and pKM1 respectively, are essentially identical except for the 9-bp repeat sequence. Comparison of their function in vivo (Table 1) indicates that the wild-type fragment has no promoter activity, whereas P_{c17} functions quite efficiently. Further comparisons indicate that P_{c17} functions about three times better than P_{482} and is nearly as efficient as our typical bacterial promoters (Table 1). Both P_{c17} and P_{482} contain perfect −10 region consensus sequences, yet differ dramatically in their efficiency. These functional differences probably reflect the adequacy of their respective −35 region recognition sequences. P_{c17} has a better −35 region sequence than P_{482}, although the two promoters also differ at other positions. Moreover, the −35 region of P_{c17} is part of a phage transcription termination signal, t_{R1} (*39*). Apparently, sequences normally used by the phage to specify termination are being used to help create the P_{c17} promoter signal. The implication of this finding is that sites of entry and exit for RNA polymerase may, in fact, share certain common features. Several other cases of overlapping promoter and terminator structures have now been reported (*32*).

Another promoter created by mutation was also isolated from a phage λ derivative, λ*cin* (*40*). This promoter, P_{cin}, is generated by a single base substitution within an intercistronic region of the phage (Fig. 3C) (*32, 39*). To examine the function of this signal in vivo we cloned identical DNA fragments from the λ*cin* derivative and the corresponding wild-type phage into the pKO vector. These two vectors, pKO-*cin* and pKM2, respectively, differ at only a single nucleotide position. Comparison of their activities indicates that promoter function is totally dependent on the single base pair alteration (Table 1). The point change that creates P_{cin} introduces an appropriate T · A pair at the most conserved position in the −10 region hexamer sequence (Fig. 3C). This creates a good (although not perfect) homology with the consensus sequence. In addition, P_{cin} has a good −35 region recognition sequence. The most efficient of the three mutant promoters, P_{cin} functions in vivo as well or better than many natural *E. coli* signals (Table 1). An interesting contradiction in P_{cin} function is the finding that P_{cin} is a rather poor promoter in vitro, notably less efficient than P_{c17} (*41*). This apparent discrepancy was explained by examining P_{cin} function in vitro on a supercoiled template rather

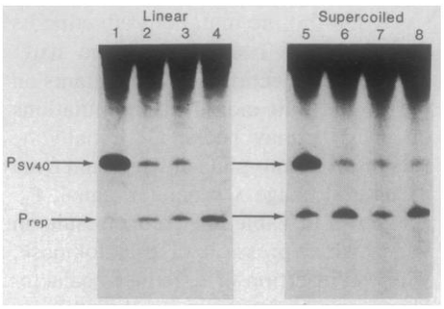

Fig. 5. Autoradiograms of polyacrylamide gel fractionation of ^{32}P-labeled RNA synthesized in vitro with the wild-type (lanes 1 and 5) and the mutant (−35 region mutants, lanes 2 and 3 and lanes 6 and 7; −10 region mutant, lanes 4 and 8) pKO-SV40 vectors as templates. Reactions were carried out with linearized (lanes 1, 2, 3, and 4) or supercoiled (lanes 5, 6, 7, and 8) templates as described (*39*). The 165-nucleotide-long transcript (P_{SV40}) initiates from the wild-type or mutant P_{SV40} promoters (see Fig. 4 and text for details); P_{rep} represents the 110-nucleotide RNA as in Fig. 2.

than on a linear template. When superhelical DNA was used, P_{cin} function was far more efficient, analogous to the situation observed in vivo (*41*). Apparently, P_{cin} belongs to that class of promoters whose functional efficiency is highly dependent on the superhelical nature of the template.

We emphasize that the three promoters characterized in these studies are of particular interest because their function, unlike that of most other promoters, is not regulated by any ancillary transcription factors. Thus, the efficiency with which these signals function depends solely on their ability to be recognized by RNA polymerase.

Promoter Mutations

The ability to select for or against *galK* expression suggests the potential of using the pKO vector system for isolating mutations in transcriptional regulatory elements. In particular, the negative selection should be useful for obtaining mutations that inactivate promoter function (that is, promoter down mutations). We chose for study a small DNA fragment (237 bp) that was derived from the genome of the mammalian virus SV40. This region of SV40 contains an efficient start site for transcription in vitro by *E. coli* RNA polymerase (*42*). The DNA sequence upstream from the start site exhibits the expected homology to naturally occurring bacterial and phage promoter signals (Fig. 4). We inserted the SV40 DNA fragment (P_{SV40}) into the pKO vector and monitored in vivo its ability to express *galK*. The results (Table 1) indicate that P_{SV40} is a relatively weak promoter, less efficient than our standard bacterial signals. The level of expression from P_{SV40}, however, is sufficient to allow the plasmid to complement a *galK*$^-$ host and also to induce galactose-dependent killing in a *gal*$^-$ background (Table 1). This typical promoter behavior makes P_{SV40} an ideal signal to demonstrate the general utility of the pKO system for selecting and characterizing promoter mutations and accurately assessing their effects on promoter function. The study of the P_{SV40} signal serves a potential dual purpose, since this region of SV40 contains the polyadenylation regulatory signal for SV40 late messenger RNA (mRNA) synthesis (Fig. 4) (*43*). Thus, mutations obtained in this region not only provide information about prokaryotic promoter function, but also represent point changes in this eukaryotic regulatory region.

The procedures developed for obtaining promoter down mutations with the pKO system are outlined below and described in detail elsewhere (*44*). Briefly, cells containing the pKO-SV40 vector were exposed to a mutagen such as hydroxylamine or nitrous acid (*45, 46*). Mutagenized plasmid DNA was then isolated and transfected into a *gal*$^-$ recipient host. These cells were plated on a tryptone broth agar (TB) containing ampicillin and galactose, and survivors were selected. These survivors represent plasmid mutants that express galactokinase at reduced levels (for example, because of mutations in *galK* or promoter down mutations). In order to select out only promoter mutations, we pooled the mutants into groups and isolated the promoter fragments from the mixture of mutated plasmids. The fragments were purified by polyacrylamide gel electrophoresis, cloned into a new pKO-SV40 vector, and subjected to the same negative selection as described above. Survivors were obtained, and plasmids from individual colonies were examined by size and restriction. Those containing the appropriately sized SV40 insert were selected for sequence analysis.

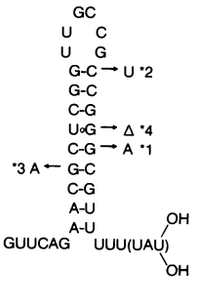

Fig. 6. The primary and potential secondary structure at the 3′ end of the t_o-terminated mRNA. Mutations shown were selected with the pKG vector system to affect termination function. Mutations 1*, 2*, and 3* were obtained by mutagenesis with nitrosoguanidine, whereas mutant 4* was obtained with benzopyrene (*54*). More than a dozen different t_o mutations have been obtained and characterized in detail with use of the pKG vector system (*48*); U, uracil.

Using the above procedure, we obtained three different point mutations within the P_{SV40} promoter region (Fig. 4). Two of these occurred in the −35 region and one in the −10 region. All three mutations affect promoter consensus sequences. The two −35 region mutations both reduced the efficiency of P_{SV40} in vivo by about 65 percent. Surprisingly, the −10 region mutant had a similar effect on promoter function in vivo. Alteration of this highly conserved −10 region position was expected to affect promoter function more severely. Indeed, when we examined the effects of these mutations in vitro, the −10 region mutation completely abolished P_{SV40} function, whereas the −35 region mutations resulted in low but clearly detectable levels of transcription (Fig. 5). Thus, there appears to be an inconsistency between the effects of these mutations on promoter function in vivo and in vitro.

This inconsistency was resolved when

Table 2. Terminators in pKG. The construction of these vectors is described in detail in (*14*). See Table 1 for information about hosts, media, and *galK* activity units. Termination efficiency in vivo was calculated from *galK* activities using P_o and P_{gal} values without terminators as 0 percent. Termination efficiency in vitro was calculated by scanning the autoradiogram shown in Fig. 7 with a Zenieth laser densitometer and determining the relative levels of terminated (*T*) and read-through (RT) RNA.

pKG Vector		*galK*$^-$ Host in indicator	*galK* Activity	Termination efficiency (%)	
Promoter	Terminator			In vivo	In vitro
P_o		Red (sick)†	2400		
P_o	t_o	Red	50	95	85
P_{gal}		Red	650		
P_{gal}	t_o	White	15	95	85
P_{gal}	t_o1*	Red	420	35	5
P_{gal}	t_o2*	Red	170	75	50
P_{gal}	t_o3*	Red	390	40	5
P_{gal}	t_o4*	Red	250	60	30

†Although host is E^+T^+, this vector produces such high levels of galactokinase that it becomes galactosemic.

the effects of these mutations were examined in vitro with supercoiled, rather than linear, templates (Fig. 5) (*44*). Supercoiling the template resulted in a marked increase in transcription from the mutant −10 region promoter (compare lanes 4 and 8 in Fig. 5). This signal now functions as well as the −35 region mutants, analogous to the results observed in vivo. Only the −10 region mutant was affected by the supercoiling. Neither the wild-type P_{SV40} nor the two −35 region mutants showed any effect. We conclude that mutation of a T · A pair to a C · G pair in this highly conserved −10 region causes promoter function to become completely dependent on supercoiling. Similarly, the function of certain natural promoters is influenced by the superhelical nature of the DNA template (*47*). Perhaps these promoters all contain −10 region sequences that are particularly sensitive to unwinding when the DNA is supercoiled. Local −10 region helix disruption should facilitate the entry of RNA polymerase into the site.

The pKO system can be used to obtain mutations in essentially any DNA sequence that functions as a prokaryotic promoter signal. Moreover, our data are consistent with the contention that deviation from consensus weakens promoter function. Most importantly, the pKO system allows us to quantitatively assess the functional effects of these mutations both in vivo (by *galK* assay; Table 1) and in vitro (by transcription assay; Fig. 5). In addition, the pKO system can be used with positive selection for *galK* expression, to obtain promoter up mutants (*48*).

Transcription Terminators

The pKO system also can be used for the isolation and characterization of transcription termination signals (*14, 15*). In this case the starting vector is a pKO derivative that contains a promoter directing *galK* expression (for example, pKG in Fig. 1B). DNA fragments are inserted between the promoter and *galK*, and their ability to reduce *galK* expression is monitored. Insertion of almost any fragment will result in some polar effect on *galK* expression. However, insertion of an authentic terminator in proper orientation results in a dramatic reduction in *galK* levels. We have inserted various terminator signals into the pKG vector and monitored their effects on *galK* expression (Table 2) (*15, 36*). The extent to which a terminator reduces *galK* activity is an accurate measure of its efficiency. This was demonstrated by measuring directly the relative amounts of RNA that stop at or that read-through the termination signal in vivo (*14, 49*). In addition, we measured the effect of varying promoter strength on the efficiency of terminator function and found that all of the terminators tested function independently of the promoters to which they are fused (*14, 36*).

Relatively few termination signals have been studied by mutational analysis, and in most cases, there is no genetic selection for obtaining such mutations. The pKG vector system circumvents this problem and allows mutational analysis of almost any terminator. In order to demonstrate the general utility of this system, we inserted a small DNA fragment that contains the phage λ terminator, t_o, into the pKG vector. Although this terminator has been studied extensively (*50–53*), no mutations affecting its function have been obtained to date. There is no selection for such mutants on the phage, and moreover, t_o mutations on phage λ may be lethal. Initially, t_o was inserted into pKG downstream from its natural phage λ promoter signal, P_o. As shown in Table 2, P_o alone results in very high expression of galactokinase, whereas insertion of t_o reduces galactokinase expression by more than 95 percent. Although t_o is quite efficient, this pKG derivative still complements a *galK*$^-$ host (Table 2). If t_o is placed downstream from a weaker promoter (for example, P_{gal}) on a pKG derivative, again it functions with greater than 95 percent efficiency. This derivative, however, fails to complement a *galK*$^-$ host and grows as a white colony on the appropriate indicator plate (Table 2). We used this derivative to select mutations in the t_o signal. Cells containing this vector were mutagenized by standard procedures (*45, 54*). Mutagenized plasmid DNA was isolated and used to transform a *galK*$^-$ host. Transformants were plated out on indicator plates and *galK*$^+$ colonies (which are red) were obtained at frequencies varying from 1×10^{-4} to 1×10^{-5}, depending on the mutagen and the mutagenic procedure. Plasmid DNA from these colonies was characterized initially by size and restriction analysis, and those retaining the appropriately sized t_o fragment were selected for DNA sequence analysis. The probability of obtaining only terminator mutations can be increased by isolating the insert from pooled groups of selected mutants and cloning it into the original pKG vector (analogous to the procedure used above for obtaining promoter down mutations).

Using these procedures, we obtained and characterized several mutations in the t_o signal (Fig. 6). These mutations all occurred within the potential stem and loop RNA structure (that is, the hyphenated-dyad symmetry structure in the DNA) immediately preceding the site of termination (Fig. 6). This stem and loop structure is a characteristic feature of almost all terminators, and various studies have indicated that its formation is required for terminator function (*32, 55*). The t_o mutations support this contention since all of them affect the stem structure and reduce its thermodynamic stability. Most importantly, the pKG vector system allows us to assess quantitatively the effects of these mutations on terminator function. Experiments carried out both in vivo (Table 2) and in vitro (Fig. 7) indicate that the t_o mutations fall into three classes. Mutants 1* and 3* have the most severe effect on t_o function.

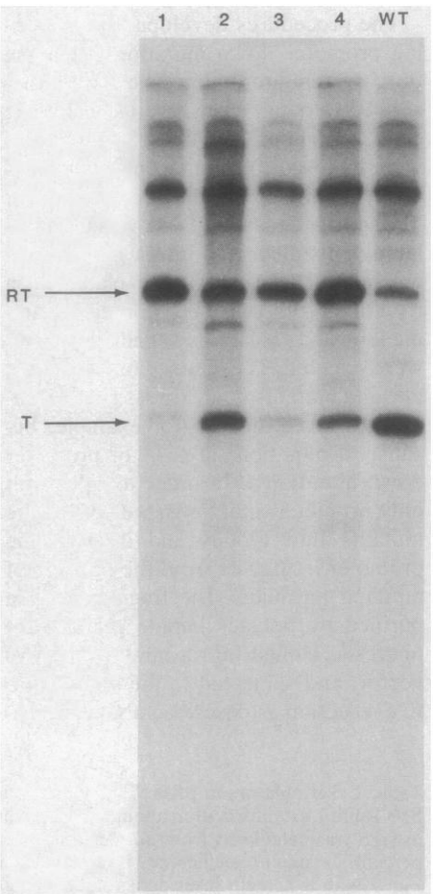

Fig. 7. Autoradiogram of a polyacrylamide gel fractionation of ^{32}P-labeled RNA synthesized in vitro from pKG vectors carrying the wild-type (*WT*) and the mutant (1*, lane 1; 2*, lane 2; 3*, lane 3; 4*, lane 4) t_o terminators (see Fig. 6). The templates were linearized at a restriction site positioned 150 nucleotides beyond t_o. In each case, transcripts initiated at the P_{gal} promoter terminate either at the t_o signal (*T*) or read-through t_o to the end of the restriction fragment (*RT*). The relative levels of T and RT RNA were quantified by scanning the autoradiogram with a Zenieth laser densitometer. Termination efficiencies were calculated and are shown in Table 2.

This is consistent with their effect of disrupting C · G base pairings in the central region of the stem structure. In contrast, mutant 2* has the mildest effect on t_o function. Again, this is consistent with its structural effect of changing a G · C pair to a G · U pair at the top of the stem. The fact that this relatively subtle change is readily detected shows the overall sensitivity of the selection system. The third class of t_o mutants, 4*, has an intermediate effect on t_o function. This deletion mutation disrupts a G · U pair and presumably results in the U residue being pushed out of the base-paired structure. Again, consistent with its functional effect, the unpaired residue has an intermediate effect on stem stability. Of course, it is also possible that these mutations affect important sequence features of the terminator signal in addition to their effects on the stem and loop structure.

The mutant terminators all function less efficiently in vitro than in vivo, although the relative order of their termination efficiencies are maintained (1* and $3^* < 4^* < 2^*$). The same is true for the wild-type t_o signal both on pKG (Table 2) and on phage λ (*53*). One possible explanation is that other transcription factors (such as rho and nus) absent in the reaction in vitro may increase the efficiency of both the normal and mutant sites in vivo. In fact, previous studies indicate that the efficiency of t_o on phage λ can be increased from 80 percent to nearly 100 percent in vitro by the addition of rho factor to the transcription reaction (*53*).

Conclusion

We have described several applications of the *galK* fusion vector system to the study of promoters and terminators recognized by the *E. coli* RNA polymerase. In particular, we have shown that the combination of both a positive and negative selection system and a simple assay allows the isolation, characterization, and mutational analysis of various transcriptional regulatory elements that otherwise could not be studied. The system lends itself to numerous other applications, including the study of various ancillary factors that affect transcription as well as characterization of the sequence-specific alterations induced by various mutagens and carcinogens (*54, 56*). In addition, the system has been recently adapted to the study of transcriptional regulatory elements in streptomyces (*57*), yeast (*27, 28*), and higher cell systems (*15, 29, 31*).

References and Notes

1. F. Cuzin, and F. Jacob, *C.R. Acad. Sci.* **258**, 1350 (1964).
2. F. Jacob, A. Ullmann, J. Monod, D. Kessler, E. Englesberg, *J. Mol. Biol.* **13**, 704 (1969).
3. D. Kessler and E. Englesberg, *J. Bacteriol.* **98**, 1159 (1969).
4. J. Beckwith, in *The Lactose Operon*, D. Zipser and J. Beckwith, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1970), p. 5.
5. J. H. Miller, W. S. Reznikoff, A. E. Silverstone, K. Ippen, E. R. Singer, J. R. Beckwith, *J. Bacteriol.* **104**, 1273 (1970).
6. W. Reznikoff and K. Thornton, *ibid.* **109**, 526 (1972).
7. D. H. Mitchell, W. S. Reznikoff, J. R. Beckwith, *J. Mol. Biol.* **93**, 331 (1975).
8. M. Casadaban, *ibid.* **104**, 541 (1976).
9. P. Bassford *et al.*, in *The Operon*, J. Miller and W. Reznikoff, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1978), p. 245.
10. M. Casadaban and S. N. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4530 (1979).
11. E. W. West, R. L. Neve, R. L. Rodriguez, *Gene* **7**, 271 (1979).
12. G. An and J. D. Friesen, *J. Bacteriol.* **140**, 440 (1979).
13. M. Casadaban and S. Cohen, *J. Mol. Biol.* **138**, 179 (1980).
14. K. McKenney, H. Shimatake, D. Court U. Schmeissner, M. Rosenberg, in *Gene Amplification and Analysis*, J. Chirikjian and T. Papas, Eds. (Elsevier/North-Holland, New York, 1981), vol. 2, p. 383.
15. M. Rosenberg, K. McKenney, D. Schumperli, in *Promoters: Structure and Function*, M. Chamberlin and R. L. Rodriguez, Eds. (Praeger, New York, 1982), p. 387.
16. D. L. Blazey and R. O. Burns, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 5011 (1982).
17. P. Youderian, S. Bouvier, M. M. Susskin, *Cell* **30**, 843 (1982).
18. G. Duester, R. M. Elford, W. M. Holmes, *ibid.*, p. 855.
19. P. Sarmientos, J. E. Sylvester, S. Contente, M. Cashel, *ibid.* **32**, 1337 (1983).
20. B. Rak, M. Lusky, M. Hable, *Nature (London)* **297**, 124 (1982).
21. H. Bedouelle, U. Schmeissner, M. Hofnung, M. Rosenberg, *J. Mol. Biol.* **161**, 519 (1982).
22. D. Drahos, G. R. Galluppi, M. Caruthers, W. Szybalski, *Gene* **18**, 343 (1982).
23. K. C. Luk and W. Szybalski, *ibid.* **17**, 247 (1982).
24. A. Miura, J. H. Krueger, S. Itoh, H. A. de Boer, M. Nomura, *Cell* **25**, 773 (1981).
25. K. C. Luk and W. Szybalski, *Virology* **125**, 403 (1983).
26. C. W. Adams and W. Hatfield, *Cell*, in press.
27. B. C. Rymond, R. S. Zitomer, D. Schumperli, M. Rosenberg, *Gene*, in press.
28. R. Zitomer, B. Rymond, D. Schumperli, M. Rosenberg, in *Gene Expression: UCLA Symposia*, D. Hamer and M. Rosenberg, Eds. (Liss, New York, in press), vol. 8.
29. D. Schumperli, B. H. Howard, M. Rosenberg, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 257 (1982).
30. P. Berg *et al.*, *Mol. Cell. Biol.* **3**, 1246 (1983).
31. H. Johansen, M. Reff, M. Rosenberg, in *Gene Expression: The Translational Step and Its Control*, B. F. C. Clark, H. Peterson, J. H. Thaysen, Eds. (Munksgaard, Copenhagen, in press).
32. M. Rosenberg and D. Court, *Annu. Rev. Genet.* **13**, 319 (1979).
33. U. Sieberlist, R. B. Simpson, W. Gilbert, *Cell* **20**, 269 (1980).
34. D. Hawley and W. R. McClure, *Nucleic Acids Res.* **11**, 2237 (1983).
35. T. Platt, *Cell* **24**, 10 (1981).
36. K. McKenney, thesis, Johns Hopkins University (1982).
37. S. Adhya, M. Gottesman, D. Court, *J. Mol. Biol.* **112**, 657 (1977).
38. B. G. Barrell, in *Procedures in Nucleic Acids Research*, G. Cantoni and D. Davies, Eds. (Harper & Row, New York, 1971), vol. 2, p. 751.
39. M. Rosenberg, D. Court, D. L. Wulff, H. Shimatake, C. Brady, *Nature (London)* **272**, 414 (1978).
40. D. Wulff, *Genetics* **82**, 401 (1976).
41. M. Rosenberg, unpublished results.
42. R. Dhar, S. M. Weissman, B. S. Zain, J. Pan, A. M. Lewis, *Nucleic Acids Res.* **1**, 595 (1974).
43. M. Fitzgerald and T. Shenk, *Cell* **24**, 251 (1981).
44. A. Chepelinsky, T. Vogel, D. Court, M. Rosenberg, in preparation.
45. J. H. Miller, in *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1972).
46. D. L. Wulff *et al.*, *J. Mol. Biol.* **138**, 209 (1980).
47. Y. Kano, T. Miyashita, H. Nakamura, K. Kurok, A. Nagata, F. Imamoto, *Gene* **13**, 173 (1981).
48. K. McKenney, in preparation.
49. M. Rosenberg and U. Schmeissner, in *Interaction of Translational and Transcriptional Controls in the Regulation of Gene Expression*, M. Grunberg-Manago and B. Safer, Eds. (Elsevier, New York, 1982), p. 1.
50. S. Hayes and W. Szybalski, *Mol. Gen. Genet.* **126**, 275 (1973).
51. M. Rosenberg, B. deCrombrugghe, R. Musso, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 717 (1976).
52. M. Rosenberg, S. Weissman, B. deCrombrugghe, *J. Biol. Chem.* **250**, 4755 (1975).
53. B. C. Howard, B. deCrombrugghe, M. Rosenberg, *Nucleic Acids Res.* **4**, 827 (1977).
54. H. Mizusawa, C. Lee, T. Kafefuda, K. McKenney, H. Shimatake, M. Rosenberg, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6817 (1981).
55. T. Platt and R. Bean, in *Prokaryotic Gene Expression*, J. R. Beckwith *et al.*, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., in press).
56. H. Mizusawa, C. Lee, T. Kakefunda, *Mutat. Res.* **82**, 47 (1981).
57. M. Brawner, M. Rosenberg, J. Auerbach, in preparation.
58. A. Levine and D. Rupp, in *Microbiology*, D. Schlessinger, Ed. (American Society for Microbiology, Washington, D.C., 1978), p. 183.

59. We thank L. Hampton and A. Venable for manuscript preparation; C. Brady and D. Sobieski for technical assistance, and M. Singer, J. Young, and J. Auerbach for helpful commentary on the manuscript.