MICROBIOLOGICAL REVIEWS, Dec. 1985, p. 398–418
0146-0749/85/040398-21$02.00/0

Vol. 49, No. 4

Uses of *lac* Fusions for the Study of Biological Problems

THOMAS J. SILHAVY[1]* AND JONATHAN R. BECKWITH[2]

Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544,[1] and Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115[2]

INTRODUCTION

The fusion of one gene to another is thought to be one of the ways in which the evolution of gene function has taken place. Many examples have now been described in which genetic techniques have been used to attach the amino-terminal portion of one protein to the carboxy-terminal portion of another while retaining the functional activity of one or both proteins (gene fusions). The generation of bi- or multifunctional proteins or the alteration of cellular location of proteins is readily imagined as the product of such gene fusions. Furthermore, in bacteria, the evolution of operons consisting of groups of genes with related functions could easily have taken place by rearrangements or transpositions that brought one set of genes under a single set of controlling elements. In these fusions (operon fusions), hybrid operons rather than hybrid genes would be generated.

Genetic fusions have from the earliest times of molecular genetics provided an important means of analyzing basic biological problems. The first example of the use of gene fusions was in the *r*II*A* and *B* genes of bacteriophage T4 of *Escherichia coli*. The *r*II*A* and *B* genes are located adjacent to one another on the T4 genome, although they are transcribed independently into separate messenger ribonucleic acid (RNA) copies. Champe and Benzer (29) isolated a T4 *r*II mutant in which a deletion (*r*1589) had removed all transcrip-

* Corresponding author.

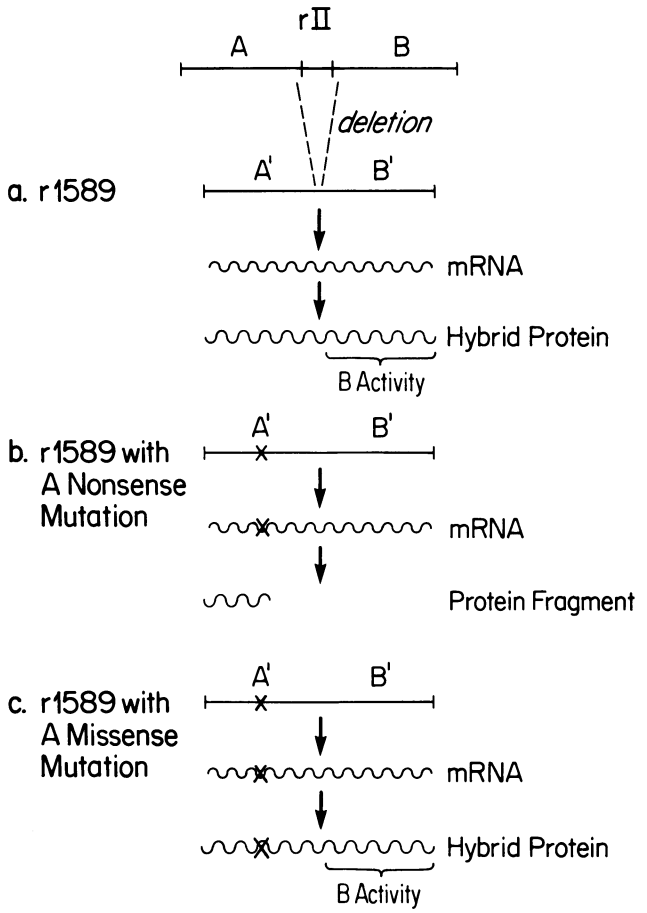

FIG. 1. Properties of the rII*AB* gene fusion (29).

tion and translation punctuation signals between the two genes and, thus, fused them into a single transcriptional and translational unit (Fig. 1). The deletion, *r*1589, which generated this fusion removed sequences coding for the carboxy terminus of the A protein and about 10% of the amino terminus of the B protein.

Despite the absence of a substantial portion of the B protein, the hybrid protein resulting from the *r*II*A-B* gene fusion still exhibited some B activity. The B activity appeared to be insensitive to missense mutations in the A portion of the hybrid protein. However, mutations which resulted in cessation of translation of the hybrid protein did, of course, lead to a loss of B activity (Fig. 1). By determining the effects of various *A* mutations on the expression of *B* activity, two very important classes of mutations were characterized. First, Crick and co-workers (31) showed that their presumed frameshift mutants in the *r*II*A* gene abolished the expression of the *r*1589 hybrid protein. Second, a class of suppressible mutants in the *r*II*A* gene when recombined into *r*1589 fusion resulted in a B^- phenotype. As a result of this and other properties of the suppressible mutants, it became clear that they represented chain-terminating or nonsense mutants (12). These findings led to very important insights into the nature of the genetic code.

These examples illustrate the basic principle of the genetic fusion approach that is used to analyze a multitude of biological problems. An assayable gene product is put under the control of another gene of interest and in this way used as a tag to follow the expression of that latter gene. To make the genetic fusion approach more generalizable, a gene that could be readily manipulated and which provided many advantages in genetic selections, screenings, and assay was needed. Beginning in the early 1960s an approach evolved using the well-characterized *lac* operon as a system for obtaining fusions.

HISTORY OF FUSIONS OF THE *lac* OPERON

The first fusions of the *lac* operon were obtained unwittingly as revertants of strong polar mutants in the *lacZ* (β-galactosidase) gene (7). It is possible to select for the functioning of the *lacY* gene (galactoside permease) independently of the *lacZ* gene. Strong polar mutants early in the *lacZ* gene eliminate both *lacZ* and *lacY* activity. Selection for revertants to growth on the carbon source melibiose (a β-galactoside), which requires the *lacY* gene product for transport but does not require β-galactosidase for its hydrolysis, yielded mutants some of which were *lacZ* $lacY^+$. These mutants were due mainly to deletions that removed the controlling elements of the *lac* operon and fused the *lacZ* gene to unknown neighboring genes. The *lacY* gene had now lost its response to the normal regulation of the *lac* operon and was presumably under the control of a nearby promoter. The properties of these fusions helped to localize the controlling elements of the *lac* operon. In 1965, Jacob and co-workers (60) described a similar approach, but this time, they devised the selection in such a way as to isolate fusions of the *lacY* gene to a nearby gene with known function, the *purE* gene. Such fusions showed repression of the galactoside permease by the addition of purines to the growth media.

In all of the cases just described, the fusions isolated were due to deletions with one endpoint within the *lacZ* gene and the other in the foreign operon. Because of the screening used, it would not have been possible to obtain deletions in which an active hybrid β-galactosidase was formed. Somewhat later, a very different approach allowed the first isolation of operon fusions in which the intact *lacY* and *lacZ* genes were brought under the control of a foreign promoter. The *lac* operon was transposed from its normal position on the chromosome to a position close to the *trp* operon (8). Located between the two operons was a locus which allowed for the selection of operon fusions. The *tonB* gene product is required for productive infection by phages T1 and φ80 and for sensitivity to certain colicins. Selection for resistance to these agents often yields deletions, certain of which also extend into neighboring genes. With the *lac* and *trp* operons arranged in the proper orientation, selection for *tonB* mutants yielded at low frequency deletions which brought the *lac* operon under the control of the *trp* repressor (76). Using appropriate starting strains, it was possible to obtain deletion strains in which the entire *lac* operon was under tryptophan control (91).

HISTORY OF *lacZ* GENE FUSIONS

It was not until the early 1970s that it became clear that β-galactosidase had features which allowed it to be manipulated in the same way as the *r*II*B* gene. First, it was shown that a substantial portion of the amino terminus (amino acids 11 to 41) of the enzyme could be removed without eliminating the potential of enzymatic activity (1). A deletion, M15, missing this region produced a protein which could still be activated by antibody to give β-galactosidase activity. Next, Muller-Hill and Kania (81) showed that genetic fusions could

be obtained in which a β-galactosidase missing its amino terminus was still active. As with the first *lac* operon fusions, the starting strain carried a strong polar mutation (*lacZU118*, an ochre mutation) in the *lacZ* gene. This mutation, which mapped early in the gene, was subsequently shown to alter a codon corresponding to amino acid 17 of the protein (127). Given that this early region of the protein was not essential for enzymatic activity, it seemed possible that removal of the early part of the gene up to or beyond the site of the mutation and fusion to a neighboring gene might yield a hybrid protein with β-galactosidase activity.

Lac^+ derivatives of the polar mutant strain were isolated. Several classes of revertants were found: mutations within the ochre codon which restored translation of the protein, ochre suppressor mutations, and deletions which removed the ochre codon and extended into the nearby *lacI* gene, thus restoring translation of the protein. These genetic fusions were shown directly to code for hybrid proteins the amino terminus of which corresponded to a portion of the *lac* repressor and the carboxy terminus of which corresponded to β-galactosidase. In some cases, the hybrid proteins exhibited both β-galactosidase and *lac* repressor activity! These latter fusions illustrate a feature of structural genes for proteins which may be more common than was imagined. There are now other examples of genes in which substantial amounts of either the 5′ end or the 3′ end (or both) of the gene can be deleted and the enzymatic activity of the protein still remains (15, 57a, 78, 81, 117; see below). This allows potentially for broad extension of the gene fusion technique to other systems.

In all the cases described to this point, genetic fusions were isolated between two genes which were normally located close to each other on a phage or bacterial chromosome. Or, in the case of the *trp-lac* fusions, the genes were brought close to each other by approaches which were not readily generalizable. To make the genetic fusion approach with the *lac* operon more useful, it was necessary to devise techniques which would allow the fusion of the *lac* operon to any gene or operon. Beginning in 1976, a series of papers described first in vivo and then in vitro methods for fusing the *lac* operon to a wide variety of genes including cloned eucaryotic genes.

GENERAL TECHNIQUES FOR CONSTRUCTING *lac* FUSIONS

In developing a general method for constructing *lac* fusions, two important problems must be solved. First, the *lac* genes must be moved to a chromosomal site that is closely linked to the target gene, and they must be positioned in the same transcriptional orientation. Next, a genetic rearrangement that creates the appropriate novel fusion joint must be selected. To illustrate these problems, let us consider the *trp-lac* fusions described above. In this case, the *lac* genes were moved to a site adjacent to *trp* by using φ80 *lac* transducing phages. These phages integrate into the chromosome at the closely linked φ80 attachment site. By choosing the appropriate transducing phage, it was possible to position the *lac* genes correctly. To create the novel fusion joint, deletions were selected that removed a conditionally lethal intervening gene, *tonB*. Strains carrying deletions that fused *trp* and *lac* were then identified by using an appropriate screen. Although not applicable generally, these studies provided a model on which future methods for constructing *lac* fusions were based.

A Two-Step Method for Constructing *lac* Fusions using Bacteriophage Mu.

The discovery of transposable genetic elements in the early 1970s provided the means to move the *lac* genes to any position on the chromosome. Because these elements can insert randomly causing defined null mutations (insertions) and because many of these elements carry selectable genetic markers, they have become important tools for the geneticist (63). More relevant to the present discussion is the fact that these elements provide rather large segments of portable genetic homology. One such element is the bacteriophage Mu. This temperate phage forms stable lysogens by transposition. Casadaban (22) utilized this phage to provide essentially random sites for the integration of *lac* DNA by homologous recombination.

To take advantage of the portable genetic homology provided by phage Mu, Casadaban constructed a specialized λ transducing phage, λ *p*1(209) (Fig. 2). This plaque-forming phage carries a portion of the *c* end of Mu and the *lacZ* and *lacY* genes without a functional promoter. The Mu and *lac* DNA is located in the *b* region of λ, and consequently, this phage is deleted for the attachment site. Because of this, it can form stable lysogens only be recombination with homologous sequences present in the bacterial chromosome, i.e., *lac* or Mu. These two bacteriophages, Mu and λ *p*1(209), provide the tools needed for the first general method for constructing *lac* fusions to any target gene. The scheme for these constructions is outlined in Fig. 2.

The first step in the scheme shown in Fig. 2 is the insertion of Mu into the target gene of interest. This is accomplished by selecting lysogens that have lost target gene function. Provided that the starting strain carries a deletion of the chromosomal *lac* genes, then integration of the λ *p*1(209) phage will occur primarily by recombination at Mu homology. Such an event serves to transpose the *lac* genes to the region of interest. By choosing the appropriate Mu lysogen, the *lac* genes can be correctly aligned.

The next step involves selecting the strains that have suffered a deletion event that generates the desired fusion. Two features of this system facilitate this selection. First, the Mu phage employed contains a temperature-sensitive repressor gene. Thus, at the nonpermissive temperature, the prophage is induced leading to cell death. Notice in Fig. 2 that this makes the Mu phage a conditionally lethal intervening genetic marker. Second, the *lac* genes do not have a functional promoter. Accordingly, expression of a Lac^+ phenotype requires that these genes be fused to a functional transcription start site. Thus, by selecting simultaneously for temperature resistance and Lac^+, fusions to a particular target gene can be isolated.

By incorporating the *lacZ* nonsense mutation *U118* into the λ *p*1(209) transducing phage, Casadaban was able to extend the results of Muller-Hill and Kania described above. Using the two-step method with this phage requires that the deletion which allows expression of a Lac^+ phenotype extend into *lacZ* sequences beyond the nonsense mutation, thus resulting in the formation of a hybrid gene that specifies a hybrid protein.

Despite the obvious usefulness of the two-step method for constructing *lac* fusions, it was not extensively used. This probably reflects the relative complexity of this procedure.

Mu d*lac* 1 and 2

The two-step method described above for the construction of *lac* fusions has been simplified by using a different method

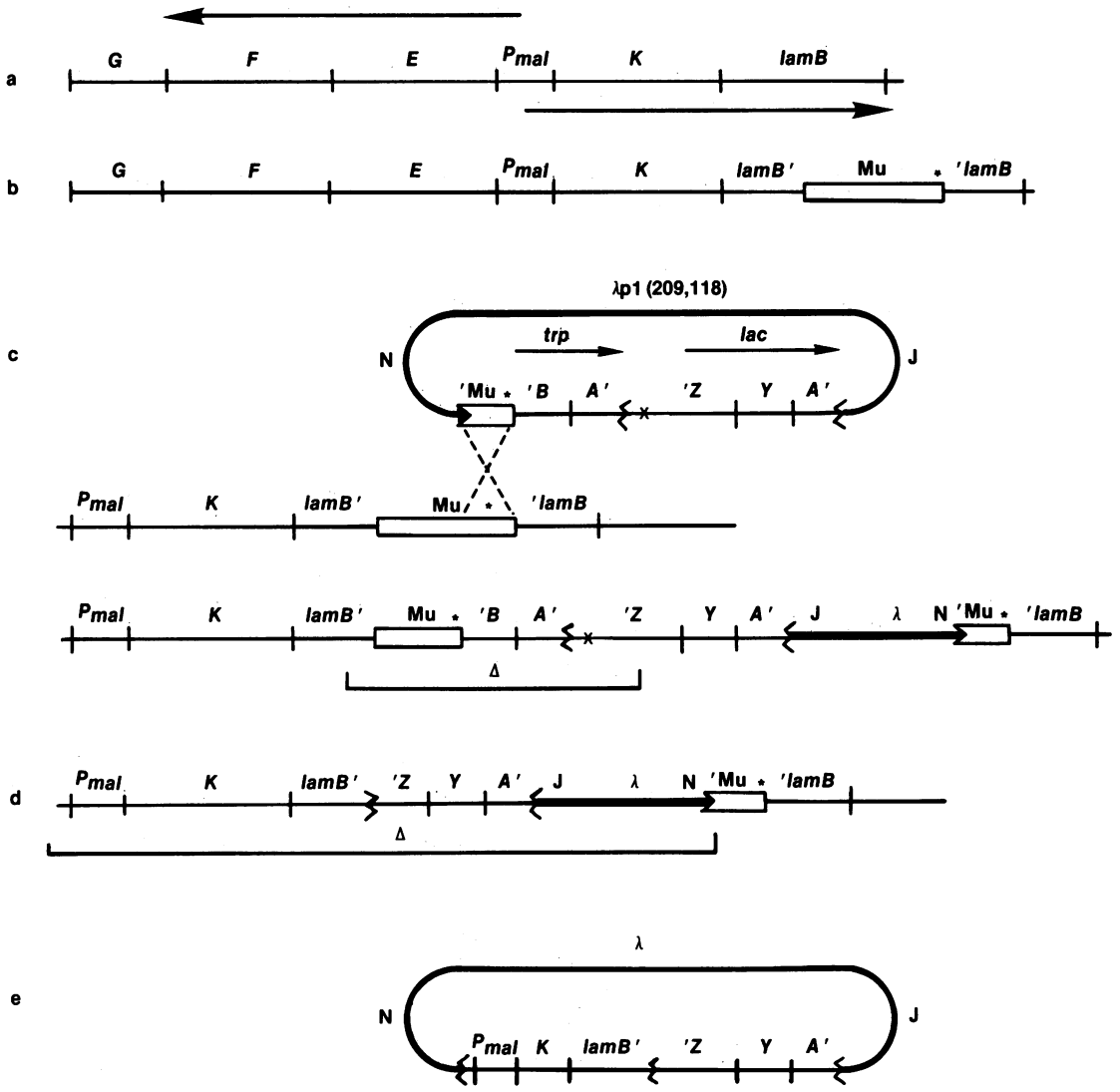

FIG. 2. Scheme for isolating *lamB-lacZ* protein fusions (36). (a) The two divergent operons that comprise the *malB* locus. Arrows indicate direction of transcription. The promoter here is designed P_{mal}. (b) A Mu *c*ts phage is inserted in *lamB* (see text for details). The phage can be inserted in either orientation; however, only those inserted with the immunity end (designated by *) promoter distal (− orientation) will yield fusions with the λ *p*1(209, 118) phage used. (c) The Mu lysogen is then lysogenized with λ *p*1(209, 118). This phage was derived from a *trp-lac* fusion (W209), and as such, it carries some *trp* DNA. It does not contain an intact *lac* promoter, and, in addition, it carries the early *lacZ* nonsense mutation *U118* (17). The parental strain, MC4100, contains a *lac* deletion, and the λ *p*1(209, 118) phage does not contain λ *att*. Therefore, lysogenization occurs primarily by recombination between the homologous Mu DNA. Since the Mu *c*ts prophage is temperature inducible, selection for Lac$^+$ survivors at 42°C in nearly all cases will yield fusions of the *lacZ* gene to P_{mal}. The purpose of the early *lacZ* nonsense mutation is to select for deletions that extend into the *lacZ* gene, producing a gene that codes for a hybrid protein (d). Deletions that do not extend past the nonsense mutation will remain Lac$^-$. The illegitimate recombination event shown below d will result in the formation of a λ specialized transducing phage (e).

for selecting the genetic event that creates the novel fusion joint. Rather than relying on a deletion event, Casadaban and Cohen (26) constructed a defective (non-plaque-forming) Mu phage (Mu d) that generates fusions by insertion. This defective Mu phage, Mu dl(Ap *lac*), carries the *lac* genes, again without a functional promoter, and the gene specifying resistance to ampicillin, *bla* (*amp* in Fig. 3 and 4). In this phage, the *lac* and *bla* genes have been substituted for structural genes that reside very near the *s* end (Fig. 3).

Four elements are known to be required for efficient Mu transposition. These include both ends (*s* and *c*) of the phage and two genes termed *A* and *B*. The *A* gene specifies the transposase, while the *B* gene functions in Mu replication (84, 116). The *lac* and *bla* substitutions in Mu dl(Ap *lac*) do not displace any of these essential transposition functions. Thus, upon infection, this phage can still transpose and form stable lysogens. If this defective phage inserts within a given structural gene in the proper orientation, then the *lac* genes can be correctly aligned with the target gene promoter. More importantly, since the small amount of Mu *s* end material present contains no transcription terminator, operon fusions can be created immediately upon insertion. Thus, by selecting for resistance to ampicillin, loss of target gene function,

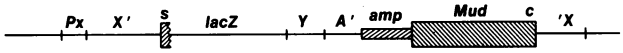

FIG. 3. Genetic structure of Mu d2(Ap *lac*) (23, 107). The ends of Mu are designated *c* and *s*, *c* representing the end on which the immunity region lies. By a number of in vivo and in vitro steps, Casadaban and Chou (23) constructed this defective (d) phage such that the *lac* genes, but not a promoter, are inserted close to the *s* end of the phage. The Mu DNA sequences remaining at the *s* end are functional for translocation; however, they do not contain transcriptional or translational stop signals. Accordingly, insertion of this phage into a target gene in the correct orientation generates a hybrid gene. If the reading frame is correct, a hybrid protein will be expressed. Thus, by using this phage, fusions can be generated in a single step.

and Lac$^+$, fusions can be created in a single step. This scheme is shown in Fig. 3.

The Mu dl(Ap *lac*) phage contains approximately 1.7 kilobases of DNA at the *s* end separating the *lac* gene from the target gene in a lysogen (69, 83). Using recombinant DNA techniques, Casadaban and Chou (23) were able to shorten this to 117 (see reference 17) base pairs without preventing transposition of the defective phage. By doing this, they were able to align *lacZ* to an open translational reading frame that extends through the end of Mu DNA. Accordingly, this defective phage, Mu d2(Ap *lac*), can be used in the manner described in Fig. 3 to generate gene fusions that specify a hybrid protein in a single step. All that is required is that the phage insert within the target gene in the correct transcriptional orientation and in the proper translational reading frame.

These defective phage, Mu dl(Ap *lac*) and Mu d2(Ap *lac*), provide a simple and effective means to generate *lac* fusions, and they have been used extensively.

The λ *plac* Mu Phages

Although the Mu d(Ap *lac*) phages dramatically simplified the construction of *lac* fusions, their use does present some rather serious problems. First, the Mu d phage is competent for transposition. Accordingly, secondary insertions occur at relatively high frequency. Stated simply, these fusions are unstable, and this complicates genetic analysis and often precludes the isolation of mutants in which fusion expression is altered. In addition, the Mu d phage, unlike the λ phage that ultimately lies adjacent to the fusion in the chromosome when the two-step method described in Fig. 2 is used, provides no obvious benefits. For example, specialized transducing phage that carry the fusion cannot be isolated, and this makes fine-structure genetic analysis and cloning more difficult (see below). Indeed, the Mu d phage is detrimental because contains a temperature-sensitive repressor and thus it renders fusion strains temperature sensitive for growth.

Komeda and Iino (65) devised a scheme for converting the Mu d prophage to λ by homologous recombination with the λ *p*1(209) phage described in Fig. 2. (This scheme is outlined in Fig. 4.) The product of this double-reciprocal recombination is essentially the same as that generated with the original two-step method. It differs only in that sequences from the *s* end of Mu are always present at the novel fusion joint. Despite the fact that this method solves the problems associated with Mu d fusions, it has not been used extensively, probably again because of its relative complexity.

The method of Komedo and Iino for converting the Mu sequences present in a Mu d*lac* fusion to λ results in the formation of a λ prophage that is flanked by the two ends of Mu (Fig. 4). The two Mu ends, *s* and *c*, are positioned in the same manner as in a normal Mu lysogen, and they contain all of the *cis*-acting sites required for transposition. Bremer et al. (17) demonstrated that this λ phage could transpose if

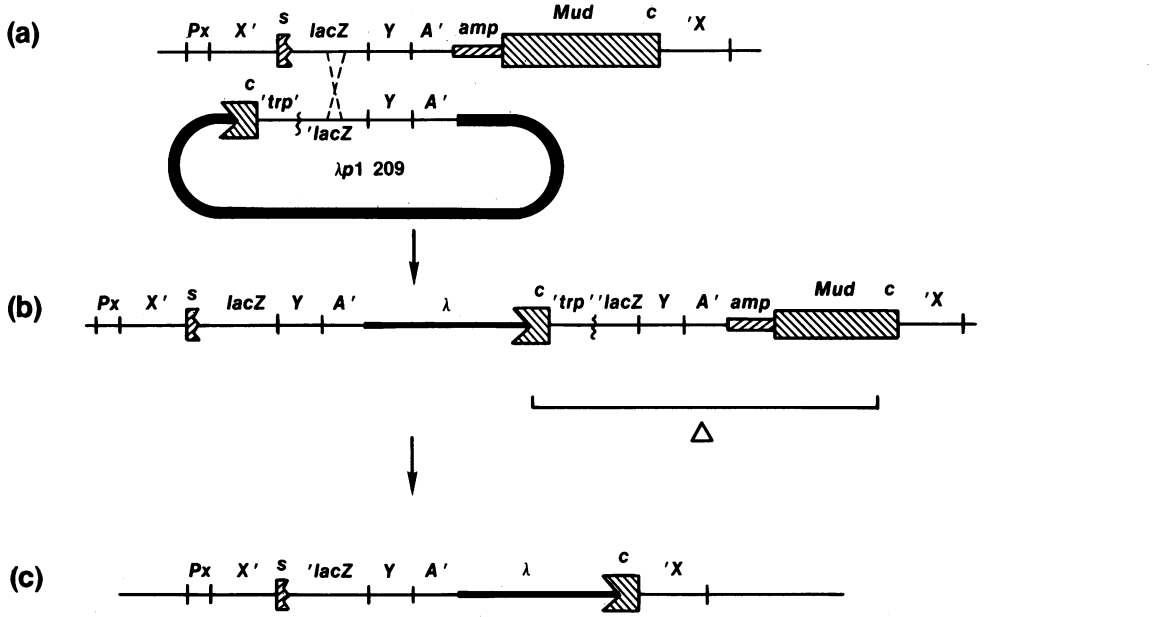

FIG. 4. Conversion of Mu d lysogen to a λ lysogen by recombination (65, 107). (a) Mu d lysogen (Ap^r, temperature sensitive) infected with λ *p*1(209) must integrate by *lacZ* or Mu homology, since this phage has no attachment site. Integration is depicted here as occurring via *lacZ* homology. (b) The double lysogen (Ap^r ts λ$^+$) loses the Mu d phage by recombination in Mu d homology (*c* end). (c) The final λ lysogen (Ap^s λ$^+$) can be selected by ability to grow at 42°C, since the Mu d prophage has a temperature-sensitive repressor. In addition to the method pictured here, integration of λ may occur by Mu homology, and excision of the Mu d prophage may occur by *lacZ* homology.

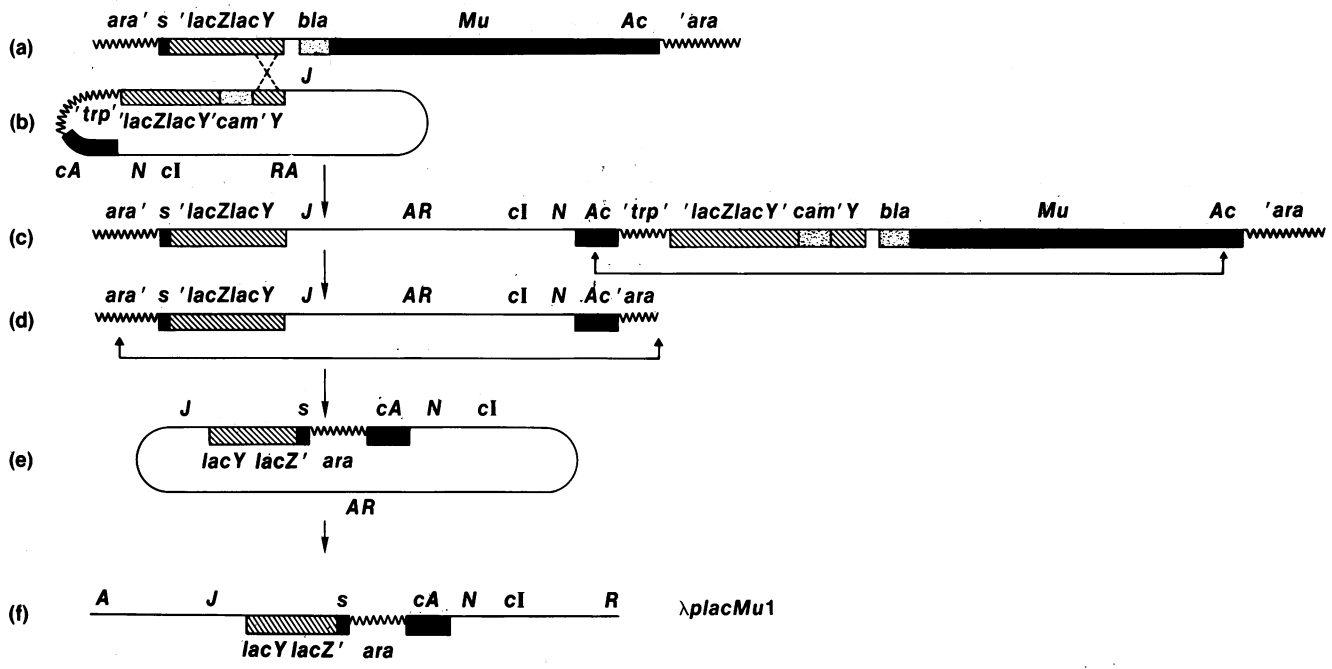

FIG. 5. Construction of λ p*lac* Mu 1. Strain GE2085, containing a Lac⁻ insertion of Mu d2(Ap *lac*)301 in the *ara* locus (a), was lysogenized with λ pSG1 (b) (17). The integration of λ pSG1 shown in the figure occurred through *lacY* homology leading to the structure shown in part c. Excision of the Mu d2(Ap *lac*)301 prophage by a homologous recombinant event involving DNA sequences in the Mu *c* region (c) led to a Lac⁻ λ lysogen, strain BRE1000, whose λ prophage is flanked by both Mu attachment sites (d). Upon UV induction, the λ prophage was excised from the chromosome (d and e), resulting in a Lac⁻ λ-Mu *lac* hybrid phage (f), λ p*lac* Mu 1. Bacterial DNA is shown by a wavy line, Mu sequences are represented by black boxes, λ DNA is denoted by a thin line, and *lacZ* and *lacY* sequences are shown by a striped box, and the shadowed areas symbolize the β-lactamase (*bla*) gene and the Tn9-encoded chloramphenicol transacetylase gene (*cam*). A prime denotes the fact that a particular gene is not completely present or is interrupted by other DNA sequences. The right Mu attachment site is designated *s*, and the left Mu attachment site is designated *c*.

the appropriate Mu functions were provided in *trans*. In addition, they isolated a specialized λ transducing phage, λ *plac* Mu, after an aberrant excision event (Fig. 5) that can integrate into the chromosome by using the Mu transposition system. More importantly, because this phage also carries the *lac* genes positioned next to the Mu *s* end as in the Mu d phages, it can create fusions in a single step upon transposition.

In practice, the λ *plac* Mu phage is easy to use. One simply infects a *lac* deletion strain with this phage and an equal number of helper phage. The helper, another λ specialized transducing phage, carries the Mu *A* and *B* genes to provide transposition functions. By selecting for lysogens that are Lac⁺ and have lost target gene function, fusions can be obtained in a single step (Fig. 6).

Recently, several derivatives of the λ *plac* Mu phage have been constructed (16). These derivatives permit the isolation of either operon or gene fusions. In addition, the new derivatives carry the gene specifying resistance to kanamycin. Like the *bla* gene in the Mu d phages, this permits the isolation of insertions even if they are Lac⁻ or if the expression of the *lac* genes is too low to allow growth on lactose. The latter point is especially relevant because by adding the sensitive indicator 5-bromo-3-chloroindoyl-β- D-galactoside (XG) to the selective media, fusions that express very low levels of LacZ activity can be detected.

A final advantage of the λ *plac* Mu phages is that they, unlike the Mu d phages, can be used to isolate fusions to genes that have been cloned onto small, high-copy-number plasmids such as pBR322. A simple protocol has been developed to select for λ p*lac* Mu insertions into the target plasmid (17). Thus, even if the cloned gene expresses no scorable or selectable function in *E. coli*, this phage can be used to isolate fusions. This extends the use of *lac* fusions to heterologous genes and organisms other than *E. coli*.

Transposon Vectors in *Bacillus subtilis* and *Myxococcus xanthus*

Although most of the methods described above for the construction of *lac* fusions in vivo cannot be adapted directly for use with other organisms, the genetic concepts should be applicable generally. Much of the technology of fusion construction utilizes transposable elements, and these are known to exist throughout biology. If certain transposons can be exploited for fusion generation in *E. coli*, then we see no reasons why this should not be true with other systems as well. Indeed, Youngman et al. (124) have adapted the transposon Tn*917* for the construction of *lac* fusions in *Bacillus subtilis*, using the same logic as was used with the Mu d and λ *plac* Mu phages. A clever selection has been devised for identifying *lac* fusions to genes that are expressed only during sporulation, and these fusions will provide important tools for the genetic analysis of this developmental pathway. Similarly, Kroos and Kaiser (66) have constructed a Tn*5* derivative that creates *lac* fusions upon insertion. Here again, fusions can be used to address the molecular mechanisms of development.

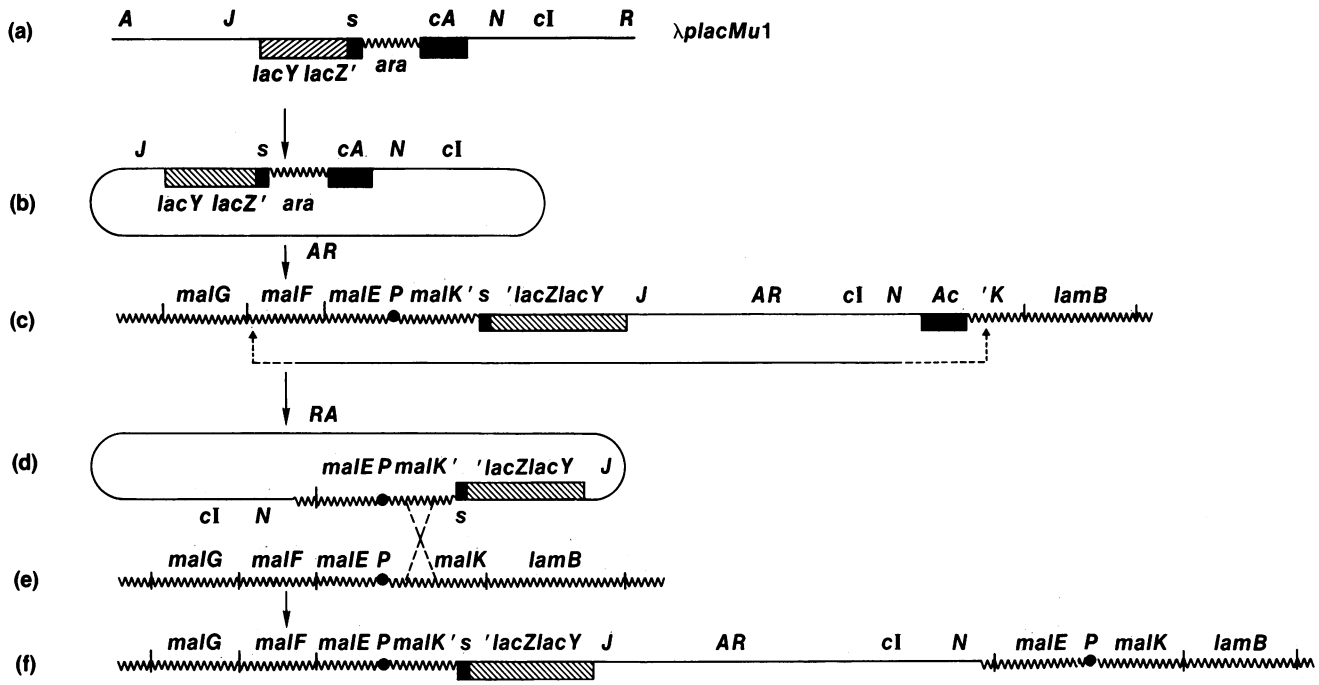

FIG. 6. Insertion of λ p*lac* Mu 1 into *malK* and isolation of transducing phages (17). Upon introduction of the mature λ p*lac* Mu 1 chromosome (a) into the cell, it circularizes (b) and inserts into the *malK* gene (c), resulting in a Lac$^+$ Mal$^-$ *malK-lacZ* gene fusion. Upon ultraviolet light induction, the prophage is excised by an illegitimate recombination event (d), leading to a Lac$^+$ *malK-lacZ* specialized transducing phage which also carries the *malE* gene. Infection of this specialized transducing phage leads to Mal$^+$ Lac$^+$ lysogens formed by integration into the *malB* region by a homologous recombination event (d to f). The black dot represents the divergent promoters (*P*) of the *malB* region.

Plasmid Vectors

The advent of recombinant DNA technology provided a powerful methodology for the construction of novel DNA joints in vitro. Not surprisingly, this technology has been extensively used for the construction of a variety of fusions. The basic principal here is the same as described above for the in vivo construction of fusions, namely, to devise a procedure in which creation of the desired novel joint causes activation of a particular enzymatic activity. With respect to *lac*, certain key modifications of the DNA sequence were required to make the system generally applicable. These modifications permitted the physical separation of the *lacZ* sequences required to specify a functional enzyme from the signals that direct transcription and translation initiation.

In the wild-type *lac* operon, *lacZ* coding sequences lie immediately adjacent to the signals for transcription and translation initiation. To separate the essential *lacZ* coding sequences from these expression signals in vitro required the presence of a useful restriction enzyme cleavage site very early in *lacZ*, i.e., before codon 26 (see below). Unfortunately, nature provided none. This problem was solved in three different ways. Gronenborn and Messing (46) introduced an *Eco*RI site at codon 5, and Casadaban et al. (24) put a *Bam*HI site at codon 8. Guarente et al. (49) took advantage of one of the *lacI-lacZ* fusions of Muller-Hill and Kania (81) and placed a variety of sites in the *lacI* portion of the hybrid gene at a position corresponding to the naturally occurring *Pvu*II site. Although many times modified, most of the currently used vectors for generating *lac* fusions in vitro are derived by using one of these changes.

The standard vector for constructing *lacZ* fusions in vitro contains a *lacZ* gene truncated at the 5′ end at one of the sites described above. Consequently, the gene is missing a promoter, a ribosome-binding site, and an ATG initiation codon. At the point of the truncation, synthetic DNA sequences containing multiple restriction enzyme cleavage sites have been inserted. Thus, if a 5′ coding sequence containing the required expression signal is cloned into one of the restriction enzyme cleavage sites so that transcription and translation are restored across *lacZ*, then a hybrid protein with β-galactosidase activity is produced. This activation of LacZ activity is usually detected with XG. Two vectors that can be used for the insertional activation of *lacZ*, pMLB1034 and pRS414, are shown in Fig. 7 and 8, respectively.

The vectors described above are used for the construction of gene fusions that specify hybrid proteins. Vectors that can be used for the construction of operon fusions also exist (27), but they are much less popular for several reasons. First, many of the investigators who use these vectors work with heterologous systems, and the translation initiation signals of *lacZ* are not recognized by the organism being studied. Second, and perhaps more importantly, plasmid vectors designed for the construction of operon fusions are difficult to use. This is because the background expression of an intact *lacZ* gene on a multicopy plasmid, even in the absence of a recognizable promoter, is quite high. Accordingly, it is often difficult to distinguish strains carrying the desired fusion from those carrying the parent plasmid. Recently, R. Simons and N. Kleckner (manuscript in preparation) have solved this problem by introducing transcription termination signals upstream of *lacZ* (Fig. 8). These vectors are likely to see extensive use.

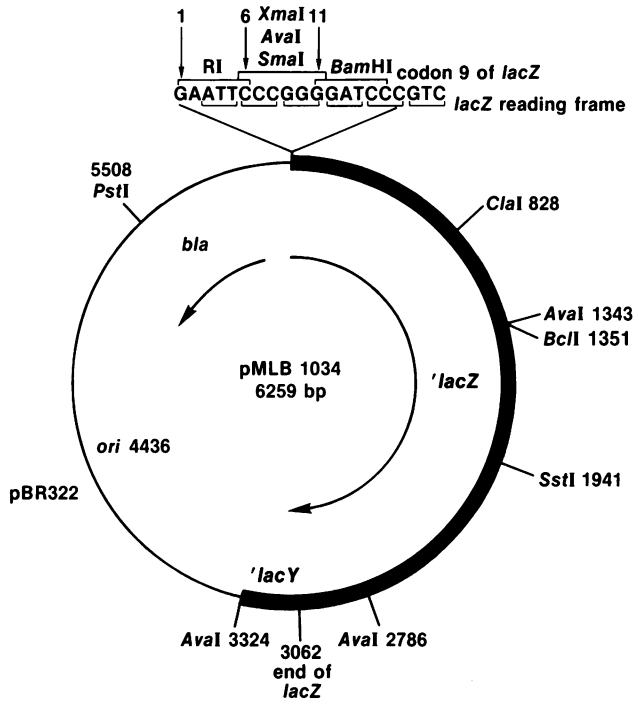

FIG. 7. Structure of pMLB1034 (107). Symbols: →, direction of transcription of the β-lactamase gene (*bla*) and *'lacZ*; ■, DNA from the *lac* operon; —, sequences from pBR322.

Selections with Plasmids

The vectors described in the previous section for the construction of *lac* fusions in vitro are most useful if the deoxyribonucleic acid (DNA) sequence of the target gene is known. This facilitates the choice of the appropriate restriction enzyme to produce a fragment that will be aligned correctly with the *lacZ* translational reading frame. If the DNA sequence of the target gene is not known, use of such vectors is hit or miss. While steps to circumvent this problem can be taken, they are tedious and time-consuming.

Berman and Jackson (14) have designed a procedure for isolating *lacZ* fusions to genes cloned on multicopy plasmids that can be used effectively even if the DNA sequence of the target gene is unknown. This procedure uses recombinant DNA techniques to position the *lacZ* gene close to and in the same orientation as the target gene, but it relies on the power of selection to identify spontaneous deletions that create the appropriate novel joint. With this method, the target gene is cloned into a plasmid such as pMLB1034 (Fig. 7) at one of the multiple cloning sites at the 5′ end of *lacZ*. The size of the DNA fragment that is inserted into the plasmid can be quite large; however, for the procedure to work, the target gene must be in the correct orientation. The resulting plasmid which contains the intact target gene upstream from the truncated *lacZ* gene is then introduced into a strain that constitutively expresses *lacY*. By selecting Lac^+, deletion events that fuse *lacZ* to an upstream gene can be identified. After retransformation to separate mutant and parent plasmids, fusions to the target gene can be identified by scoring for loss of target gene function.

This method is greatly facilitated by the use of indicator media such as lactose-MacConkey agar. As described below, this medium can be used for the selection of rare Lac^+ colonies growing out of a lawn of Lac^- bacteria. Accordingly, even though deletions that generate fusions are very rare (<1 in 10 billion), many independent fusion strains can be isolated on a single plate as Lac^+ papillae after several days of incubation. This method also allows the isolation of numerous different gene fusions, and this, as described below, can be of particular advantage.

Although the schemes described in this and the preceding section for constructing *lacZ* on plasmids may be the method of choice in certain cases, caution must be exercised in the use of such fusions for studies of gene regulation. Physiological studies may yield misleading or erroneous results owing to the high copy number of the plasmid. Before using such a construct for studies of regulation, it should be moved onto a λ transducing phage so that lysogens, each containing a single copy of the fusion, can be made. Several specialized λ transducing phages designed for this purpose have been described (Simons and Kleckner, in preparation; M. Berman and R. Zagursky, personal communication). With these phages, transfer can be accomplished by homologous recombination (Fig. 8). Alternatively, the fusion can be cloned onto the general fusion acceptor phage, λNF1955 (107).

Vectors for Constructing Tripartite *lacZ* Fusions

The chief limitation of the methods described above for the construction of *lac* fusions is that expression is dependent on appropriate sequences of the target gene, i.e., translation or transcription (or both) start sites. This limitation can be serious if the target gene is expressed at very low levels or if the target gene is derived from an organism other than *E. coli*. A solution to this problem was provided by the advent of vectors that allow the construction of tripartite *lacZ* gene fusions. These vectors can be used to express any translational open reading frame (Orf) in *E. coli* and consequently, they are often referred to as Orf vectors.

Most Orf vectors are small, high-copy-number plasmids (usually derivatives of pBR322) that contain a powerful promoter, an active ribosome-binding site, and a translation initiation codon. Located immediately downstream of these signals in the vector is a large functional fragment of *lacZ* lacking its normal expression signals. These two key pieces are aligned correctly for transcription; however, they are not in the same translational reading frame. Accordingly, the vectors confer a $LacZ^-$ phenotype. If a DNA fragment that is an open reading frame is inserted between the key pieces such that translational reading frame is restored, then a three-part or tribrid protein that exhibits β-galactosidase activity is produced. The tribrid protein contains one or more amino acids derived from the vector at its amino terminus, target gene sequences in the middle, and LacZ at the carboxy terminus. A summary of commonly used Orf vectors is presented in Table 1.

To facilitate constructions, the Orf vectors carry synthetic DNA sequences at the joint between the expression signals and *lacZ* that contain multiple restriction enzyme cleavage sites. Since the number of possible cloning sites is large, it is often possible to achieve expression of the open reading frame of interest by choosing a suitable Orf vector. If DNA sequence data of the target gene are not available, then one of several procedures for generating a collection of small fragments with different ends can be employed.

GENERAL REASONS FOR USING *lac* FUSIONS

There are a variety of purposes for which *lac* fusions are used. These include the following: to study the regulation of another gene or operon, to detect genes which are subject to

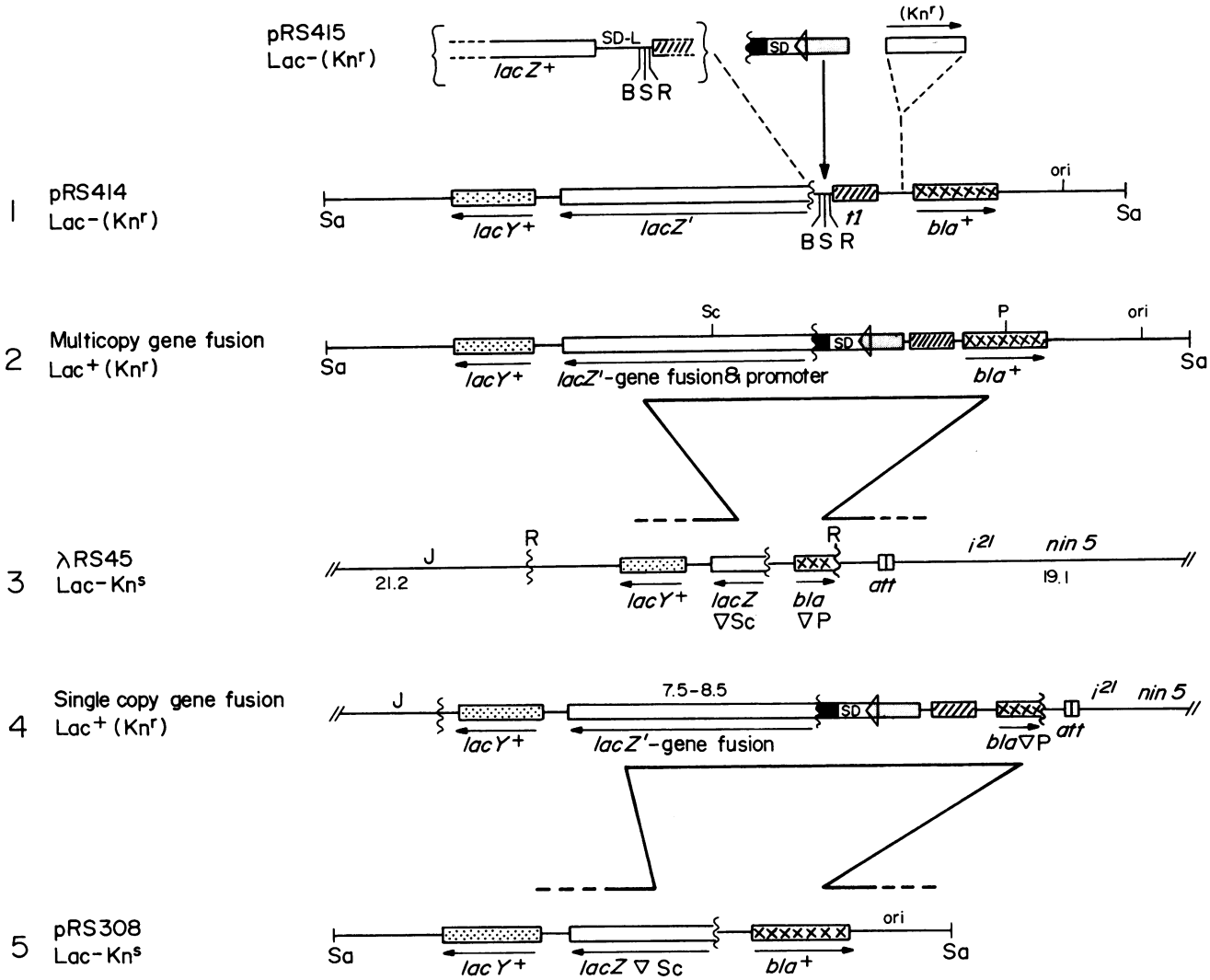

FIG. 8. Improved single and multicopy *lacZ*-based protein and operon fusion cloning tools (Simons and Kleckner, in preparation. (1) Plasmids pRS414 and pRS415 (which differ from one another as indicated in the brackets) can be used, respectively, for the construction of protein (i.e., gene) and operon fusions to the *lacZ* gene. Both plasmids contain several convenient cloning sites; in pRS414 these sites have been inserted into codon 8 of the LacZ gene (*lacZ*). In pRS415, they are situated upstream of the normal translation initiation site (SD-L) of the intact *lacZ* gene ($lacZ^+$). Both plasmids contain a strong transcription terminator (*t1*) that prevents transcription originating upstream of the cloning sites from perturbing expression of the fusion. Versions of these plasmids, called pRS414-K and pRS415-K, contain the kanamycin resistance gene (Kn^r) of Tn*903*. (2) Using pRS414, multicopy gene fusions are constructed by inserting fragments containing the promoter (⇨) and 5′ end (⊂⇨) of an exogenous gene, such that its reading frame coincides with that of *lacZ*. In a similar fashion, operon fusions are constructed by inserting promoter-bearing fragments into pRS415. The low basal level of *lacZ* expression from pRS414 (<0.1 units) and pRS415 (ca. 50 units) provides easy detection of protein and operon fusions on appropriate indicator plates. (3) The Lac^- bacteriophage λRS45 contains sequences homologous to regions that flank the fusion junction of the multicopy fusion plasmids. (4) By growing λRS45 on a transformant containing a multicopy fusion plasmid, Lac^+ recombinant transducing bacteriophage can be obtained which will contain a restored *lacZ′* (or $lacZ^+$ gene, along with the constructed fusion and the upstream *t1* terminator. These bacteriophage can subsequently be lysogenized to produce single-copy fusions. Use of Kn^r vectors in step 2 will facilitate the detection or selection (or both) of such transducing phage. (5) In an analogous fusion, the Lac^- Kn^s plasmid pRS308 can be used to cross single-copy fusions back onto a multicopy plasmid. B, *Bam*HI; S, *Sma*I; R, *Eco*RI; Sa, *Sal*I; Sc, *Sac*I; P, *Pst*I; SD, translation initation site of exogenous gene; Sc, deletion of *lacZ* retaining only the distal half of the gene; P, deletion of *bla* retaining only the proximal half of the gene. Lengths of phage arms and inserts are indicated in kilobases. Recombination arms are indicated by the heavy lines between 2 and 3, and 4 and 5.

a particular regulatory signal, to study the mechanism of localization of a protein to an intracellular compartment or to the extracellular space, to detect and assay a protein which has not been identified or for which no assay exists, to detect a gene, the protein product of which is known, but for which there is no simple method for localizing or cloning the gene.

There are several reasons why the *lac* genes have been so widely used for genetic fusions studies. First, the *lac* operon was one of the most intensively studied genetic systems, so that numerous genetic and biochemical aspects of the system were known. Second, since lactose, the substrate of the system, is a sugar which can be utilized by *E. coli* as a sole carbon source, the several indicator media available for detecting sugar metabolism could be used (107). This meant that when the *lacZ* and *lacY* genes were put under the control of a set of regulatory elements for another gene, the

TABLE 1. Orf vectors

Reference	Promoter	NH_2 terminus	COOH terminus
Koenen et al. (64)	*lac*	*lacZ*	*lacZ*
Zabeau and Stanley (125)	λ Pr	λ *cro*	*lacI-lacZ*
Gray et al. (45)	*lac*	λ cI	*lacI-lacZ*
Shultz et al. (102)	*lac*	*lacZ*	*lacZ*
Weinstock et al. (121)	*ompF*	*ompF*	*lacZ*
Weinstock and Weisemann (unpublished data)	Colicin E1	Colicin E1	*lacZ*
Weis et al. (122)	*lac*	λ *cro*	*lacZ*

expression of those controlling elements could be monitored and genetic selections carried out on these same indicator media. These media also vary in their properties so that different ones can be used for different purposes. For instance, on lactose-MacConkey agar, Lac^- bacteria form white colonies and Lac^+ bacteria form red colonies. Even very weakly Lac^+ bacteria will form colonies with detectable red color. Thus, this medium allows sensitive detection of increased levels of *lac* expression in fusions. In contrast, on lactose-tetrazolium agar, Lac^+ colonies are white, and Lac^- colonies are red. On this medium it is relatively easy to detect colonies with reduced *lac* expression, even if they are still able to grow on lactose. On both these media, Lac^+ revertants can be detected on a lawn of Lac^- bacteria. Another medium with properties similar to MacConkey is eosin-methylene blue medium. While this medium has been less extensively used, recent studies indicate that in certain situations it is more useful than the other two (34).

The third feature of the lactose system is key. The fact that lactose (glucose-1,4-β-D-galactoside) is a disaccharide provides the opportunity to generate numerous functional analogs of lactose in which the glucose moiety is substituted. The analog *o*-nitrophenyl-β-D-galactopyranoside is a substrate of β-galactosidase which upon hydrolysis by the enzyme yields *o*-nitrophenol which has a yellow color. The result is one of the most sensitive enzyme assays available. Levels of β-galactosidase activity as low as 0.001% of wild type have been detected in *E. coli lac* mutants! This means that if one is studying a system in which the gene product is not readily assayable or the assay is not very sensitive, quantitative analysis becomes possible with *lac* fusions. Further, in vitro studies on gene expression can rely on the β-galactosidase assay, if fusions have been constructed.

Three other analogs will be mentioned which can be useful for genetic studies on regulation. One of these, *o*-nitrophenyl-thiogalactopyranoside (tONPG) is transported into the cell by the product of the *lacY* gene. If so transported, the analog causes bacteriostasis. Mutants resistant to tONPG can be selected with either reduced or abolished *lac* expression. Among such mutants, which have lowered *lacY* expression, are *lacY* structural gene mutations, polar mutations in *lacZ*, and promoter mutants (58, 111). Thus, tONPG provides a selection for regulatory mutants which abolish gene expression or for leaky promoter mutants which have been important in studying genetic regulatory elements.

Phenylethylthiogalactoside is an inhibitor of β-galactosidase. If a strain is expressing the *lac* operon at an intermediate or low level such that it is still able to grow on lactose, the inclusion of the appropriate concentration of phenylethylthiogalactoside in the medium will reduce hydrolysis of lactose enough so that the bacteria are no longer able to grow on lactose. Then, selection for Lac^+ on such media can yield mutant derivatives with increased expression of such a fusion (54). This approach can be used to isolate promoter mutants with increased activity or other types of regulatory mutants.

The analog XG is a substrate of β-galactosidase which yields a blue color upon hydrolysis. It provides yet another indicator medium for examining lactose metabolism. However, in addition, it is not dependent or, at least, not totally dependent upon the *lacY* gene product for its transport into the cell. Thus, expression of fusions in which the *lacY* gene is not present can still be assayed on solid media by observing the degree of blue color when XG is present (107).

The nature of the β-galactosidase protein itself has proved useful in gene fusion studies. β-Galactosidase is a tetrameric protein made up of identical subunits. The molecular weight of the subunit, the product of the *lacZ* gene, is 116,000. From the analysis of numerous gene fusions comes the finding that up to 26 amino acids can be removed from the amino terminus and substituted for by other amino acid sequences with very little effect on the specific activity of the enzyme (18, 37). Hybrid proteins have been characterized in which the contribution to the amino terminus of β-galactosidase from the other protein is as little as 5 amino acids (95) or as much as 500 or more (see, e.g., references 39 and 109). There are some situations in which the nature of the amino terminus of the hybrid protein can have a profound effect on enzymatic activity. In the cases of certain fusions of the *lacZ* gene to the genes for cell envelope proteins, the hybrid protein is found in the membrane fraction. While many such fusions produce high levels of β-galactosidase activity, others exhibit dramatically reduced enzymatic activity as a result of the membrane localization of the protein (56, 75, 86).

Thus, hybrid proteins with β-galactosidase activity have subunit molecular weights of anywhere from approximately 114,000 to 116,000 and up. These proteins are among the largest found in the bacteria. When the pattern of *E. coli* proteins is examined on sodium dodecyl sulfate-polyacrylamide gel electrophoresis, the hybrid proteins are usually found in a region of the gel in which very few other proteins migrate. As a result, the proteins, if they are made in reasonably high levels, are always easy to visualize by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Furthermore, and more importantly, substantial purification of the proteins can be achieved by passage through a sizing gel in which the hybrid proteins are separated from most other cellular proteins. This property of gene fusions to *lacZ* has proved extremely useful as described subsequently in this review.

GENETIC MANIPULATIONS WITH FUSIONS

The isolation of strains carrying *lac* fusions provides many useful tools for genetic analysis of regulatory systems and the target gene. Many of these advantages derive from the fact that either in the construction of the fusions or by subsequent manipulations, bacteriophage λ DNA is located adjacent to the fusion. Accordingly, many of the techniques developed in the study of λ can be applied to the target system or gene (107).

In fusions constructed by certain of the in vivo approaches described in this review, the adjacent λ phage is bounded neither by its normal attachment sites nor by genetic homology. When the phage is induced to multiply, e.g., by ultraviolet light, excision can take place only by illegitimate

recombination. As a result, the population of phage found in the lysate induced represents a collection of phage carrying differing amounts of adjacent genetic material. Thus, by employing an appropriate screen, transducing phage carrying the fusion are readily obtainable (Fig. 2 and 6). These phage provide a means of moving a fusion from one strain to another. In this way, the effects of different regulatory mutations on the expression of the fusion can be tested by transferring the transducing phage to the appropriate strain background. In addition, mutants in a structural gene under study or in its controlling elements can be transferred by recombination from the transducing phage to a chromosomal copy of a gene or vice versa. For instance, promoter mutants obtained by use of a *lac* fusion can be recombined onto a chromosomal copy of a gene, allowing study of the effects of the mutation on the native gene (13).

Besides permitting the isolation of transducing phage, the illegitimate excision of the adjacent λ phage provides the opportunity to obtain a set of phage that allow fine-structure mapping of nearby genes. That is, as induction of a normal lysogen yields a set of λ *gal* phage carrying deletions of differing extents of the *gal* operon (19), the fusion strains yield a similar set of phage for genes located next to the fusion. This approach was essential, for instance, for the fine-structure mapping of the *malB* locus (108) and the *phoA* gene (96).

In addition, λ phage lysates from fusion strains will often include phage that carry nearby genes unrelated to the fused gene. These phage can then be used to study the other gene or as a source of DNA for subsequent subcloning onto plasmids (96).

The converse of using induction to obtain λ transducing phage is to select for bacteria that survive the inducing event. In this manner, deletions that remove the lethal prophage and which extend to various extents into the adjacent genetic material can be identified. This technique was first used with a λ phage integrated at the normal attachment site or one of a variety of secondary attachment sites (99, 101). With fusions, this technique can now be used for any region of the chromosome. A method for simplifying this selection has been described and used for the isolation of strains deleted for the *ompB* locus (40).

It is important to note that mutations that alter the hybrid gene or that affect its expression can be identified on the λ transducing phage itself. This can be done by using an appropriate selection or by screening plaque color on indicator agar. Moreover, these phage are a convenient substrate for local mutagenesis. By growing the phage on a mutator strain or treating it directly with various mutagens, mutations that alter the target gene or its controlling elements can be obtained. In this manner, the mutagenic treatment can be focused to a particular region (107). Several examples of the successful use of this approach are described in a following section.

USES OF *lac* FUSIONS TO STUDY THE REGULATION OF GENE EXPRESSION

Definition and Study of Regulatory Genes

An important part of determining the mechanism of regulation of a gene or set of genes is characterizing regulatory genes for the system. The ability to isolate constitutive mutants for the *lac* operon, for instance, led to the discovery of the *lacI* gene and the formulation of the repressor hypothesis. However, in many systems the approaches are not available for the detection of such mutants. In those cases, putting the *lac* operon under the control of the particular regulatory system often allows ready isolation of regulatory mutants.

Detection of New Regulatory Genes

Chou et al. (30) used the genetic fusion approach to detect a gene which regulates transposition of Tn*3*. They fused *lac* to the structural gene (*tnpA*) for the Tn*3* transposase, the enzyme responsible for promoting transposition. Ordinarily, this gene is expressed at a very low level. The resulting low level of *lac* expression of the *tnpA-lacZ* fusion was not sufficient to allow growth on lactose as the carbon source. However, from the fusion strains, Lac^+ derivatives were isolated which contained mutations in a gene, *tnpR*, coding for a repressor of the transposase. These mutations inactivated the repressor, resulting in high levels of *lac* expression from the *tnpA* promoter and permitting growth on lactose. Thus, these findings revealed the existence of a repressor. Without the availability of the gene fusion approach, the detection and characterization of this gene would have been much more difficult. A similar approach was used to identify a regulatory gene for the *guaAB* operon (115).

Cloning Regulatory Genes

Because *lac* fusions confer an easily selectable or scorable phenotype, they provide a useful means to identify clones which carry regulatory genes from a gene bank or library. For instance, the *ompR* gene specifies a positive regulatory protein required for the expression of *ompF* and *ompC*. Thus, an *ompR ompF-lac* fusion strain is Lac^-. By selecting Lac^+ transductants, it was possible to identify those λ phage that carry *ompR* in a *E. coli* library (114). Here again, without fusions this would have been much more tedious. The generality of this approach is demonstrated by its use to clone the regulatory gene for cholera toxin from *Vibrio cholerae* (77).

Isolation of New Classes of Regulatory Mutants

The genes involved in maltose utilization in *E. coli* are subject to positive control by the product of an activator gene, *malT*. This gene was identified originally by the detection of mutants in which the gene was inactivated giving a Mal^- phenotype. However, it was difficult to develop an approach for isolating mutants which altered the *malT* gene to give a constitutive phenotype (*malT*^{c} mutants). Such mutants were important for analyzing in detail the mechanism of regulation. The advent of the *lac* fusion technology permitted a selection for such mutants. Fusions of *lac* to the *malPQ* operon were constructed. Again, the basal level of expression of this operon was too low to allow growth on lactose. Many of the Lac^+ derivatives of this strain contained *malT*^{c} mutations. The mutants were then used to show that positive control of the maltose system differed from that controlling the arabinose operon (32). Further, the *malT*^{c} mutations have been useful in other studies in which constitutive expression of the *mal* genes was required.

In some cases, the basal level of a regulated operon may be so high as to allow growth of *lac* fusion strains on lactose. In those cases, selections for Lac^+ may be done in the presence of phenylethylthiogalactoside (see above).

Characterization of the Controlling Elements of Genes or Operons

In addition to obtaining mutants in regulatory genes, the fusion approach can be used to isolate mutations in promoters, operators, and sites for interaction of positive control factors. Such mutations have been essential for studying the interaction of proteins with DNA and for elucidating control mechanisms. Yet, in many systems, there is no obvious way to isolate such mutants. Let us take the structural gene for tyrosine-transfer RNA as an example. For this particular gene, there is a selection for gene inactivation which should yield both structural gene mutations and mutations in the *tyrT* promoter. However, there is no way to distinguish the two. Fusing the *lac* genes to the *tyrT* promoter provides a new selection and also a screening procedure for detecting promoter mutants. The particular advantage of such fusions is that they include two *lac* genes, *lacZ* and *lacY*. Mutations in the *lac* structural genes will, for the most part, affect only the gene in which they lie. In contrast, mutations in the promoter to which they are attached should affect both genes in parallel. Thus, in the fusions, in contrast to the original *tyrT* gene, promoter mutants are distinguishable from structural gene mutations. (The exception to this generalization is polar mutations in the *lacZ* gene. However, by seeking only those promoter mutants which are leaky (still produce some β-galactosidase) it is possible to distinguish them from polar mutants.)

The selection used in this case was for resistance to tONPG (a selection for low levels of *lacY* expression, see above), and resistant clones were screened for leaky levels of β-galactosidase on XG medium. The mutants isolated in this way were the first promoter mutants for a stable RNA gene (13).

Other selections for promoter mutants that exploit properties of fusions have also been devised. Strains in which the *lacZ* gene is fused to a gene specifying an exported protein often exhibit a phenotype termed overproduction lethality. This phenotype is thought to be the result of the aborted attempt of the cell to export large quantities of hybrid protein containing sequences of β-galactosidase (see below). With *ompF-lacZ* fusions, a novel variation of this phenotype has been uncovered. In this case, it was found that λ transducing phage carrying the hybrid gene are unable to grow on a host harboring a regulatory mutation that causes increased *ompF* expression. This higher level of hybrid gene transcription, coupled with the increased copy number that occurs during lytic phage infection, apparently produces sufficient amounts of hybrid protein to interfere with phage growth. By selecting for mutant transducing phage that can grow on this particular host, it is possible to obtain mutations that decrease or abolish hybrid protein synthesis or export (112). Using the screening procedure outlined above for *tyrT*, promoter mutants can be identified (113).

The phenotype of overproduction lethality is not associated with *lacZ* fusions to genes that specify exported proteins only. For example, high-level expression of certain *recA-lacZ* fusions is lethal as well (123). Although the mechanism for this growth defect is not yet understood, it is possible to isolate promoter mutants by using a fusion transducing phage and a host that overproduces RecA (123a).

Fusions are particularly advantageous for the detection of controlling element mutations in those genes or operons which are part of a regulon. A regulon is a group of unlinked genes which is subject to the same regulatory protein(s). In such systems, it is often possible to obtain regulatory mutations, but they are always in the structural gene for the regulatory protein. For instance, genes in three different operons determine the ability of *E. coli* to utilize maltose as a sole carbon source. As mentioned above, *malT* mutants which affected the expression of all three operons had been readily isolated at an early stage of the study of this system. However, mutants which would affect the controlling elements of any one operon were difficult to detect, since any screening or selection system would only assay for overall Mal phenotype. However, by fusing the *lac* genes to any one of the *mal* operons, it becomes possible to isolate controlling element mutants for that locus. For instance, in the absence of maltose, the fusion strains are Lac^-, and selection for Lac^+ yields either $malT^c$ mutants or up-promoter mutants for the particular operon (9). On the other hand, *lac* fusions to the *malPQ* operon were used to isolate promoter-down mutations (53). These mutations are contributing to an understanding of how this operon is positively controlled.

Distinguishing Transcriptional from Posttranscriptional Control

When a gene or operon is subject to a regulatory mechanism, it is often difficult to determine whether the control is operating at the transcriptional or the translational level. The possibility of constructing both operon and gene fusions of the *lac* operon to other genes permits a distinction between the two levels of control. In an operon fusion, the translational signals of the *lacZ* gene are intact, and the expression of the *lac* genes is dependent only on the transcriptional signals of the gene to which it is attached. In contrast, in gene fusions, both the transcriptional and translational signals of the *lacZ* gene have been replaced by those of the gene to which it is fused. If a regulatory mechanism is operating at transcriptional level, then the *lacZ* gene in both gene and operon fusions will be affected. However, if regulation is occurring at the initiation of translation, only the gene fusion will be affected. This approach was used to demonstrate that a gene coding for a ribosomal protein was regulated by autogenous control at the initiation of translation (79).

In the case of the synthesis of gluconate-6-phosphate dehydrogenase, posttranscriptional regulation was shown by the use of gene and operon fusions (3). Furthermore, a detailed analysis of gene fusions revealed an unusual aspect to this control mechanism. A region internal to the *gnd* gene appeared essential for this regulation to take place (4). The basis for this novel mechanism is currently being investigated.

Recently, it has been shown that translation of a particular messenger RNA can be inhibited by pairing with a complementary (antisense) RNA sequence. In several cases, e.g., *ompF* or the transposase of Tn*10*, this form of control appears to be important physiologically. In both of these cases, *lac* fusions were useful for demonstrating that regulation occurred at the level of translation (80, 98, 110).

One qualification to the studies described in this section should be mentioned. Under some conditions, it may well be that an extreme reduction of translation of a gene will, in turn, reduce the transcription of downstream genes. Such an effect is analogous to that seen with polar mutants in operons. Alternatively, a reduction in translation may cause dramatic destabilization of the messenger RNA leading to rapid degradation. In such a case, while the primary control was at the level of translation, the comparison of effects on

the two fusions could lead to the false conclusion that it was at the transcription level.

Demonstration of Translational Coupling

This example of the utility of the genetic fusion approach represents a rather esoteric case, but does illustrate the ways in which fusions can be used when no other obvious approach to a problem can be envisioned. Translational coupling refers to a situation in which two adjacent genes in an operon have overlapping translational termination and initiation signals. In the *trp* operon of *E. coli*, this is the case for the *trpE* and *trpD* genes, the first two genes in the operon. Oppenheim and Yanofsky (87) showed that polar mutants in the *trpE* gene had a much more drastic effect on expression of the *trpD* gene than on the subsequent downstream genes, *trpC*, *trpB*, and *trpA*. A similar overlap in punctuation signals was observed upon sequence analysis of the *trpB* and *trpA* genes. However, since there were no genes in the operon downstream to the *trpA* gene, there was no way to do experiments analogous to those done with *trpE* and *trpD* to demonstrate translational coupling. This difficulty was overcome by introducing downstream genes to *trpA* by using the *lac* fusion approach. Now, β-galactosidase could be used to compare polar effects of *trpB* mutants on *trpA* expression and downstream genes. The results indicated that these two genes were also translationally coupled (2).

Physiological Studies on Regulation

The existence of a *lacZ* fusion to a gene of interest, in many cases, allows a more detailed analysis of the factors regulating a gene than would otherwise be possible. In some cases, there is either no assay for the gene product being studied or the assay is laborious. In these cases, uses of the β-galactosidase assay overcomes the limitations of the system. In this way, the effects of mutations in other genes or of changes in physiological conditions on the expression of a particular gene can be easily monitored. Fusions of this sort have been used to study, among others, the anaerobic control of the *glpA* gene (67), the regulation by fatty acids of the *fadL* gene (94), and the factors which regulate the *sfiA* gene (59), a gene involved in cell division. In other cases, the sensitivity of the assay is insufficient for quantitative analysis. For example, fusions have been used to show that the Tn*10* transposase is present at one molecule or less per cell (E. Raleigh and N. Kleckner, Proc. Natl. Acad. Sci. USA, in press).

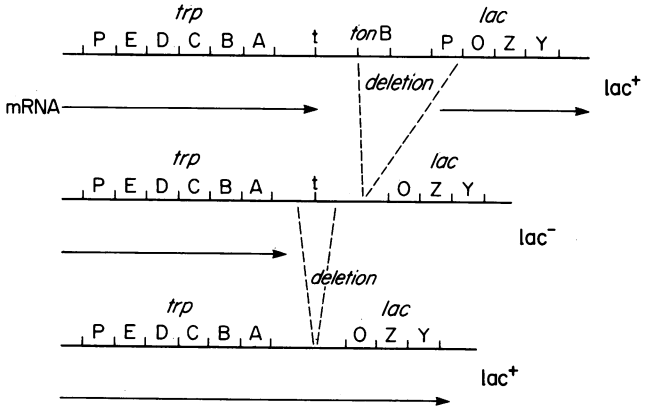

FIG. 9. Use of *lac* fusions to select mutations in a messenger RNA termination signal (48).

Studying the Regulation of a Gene in the Absence of Its Product

In many instances, it is crucial for the analysis of a regulatory mechanism to be able to study regulation when the product of the gene being regulated is not active. Since, in the formation of most *lac* fusions, a portion of the gene to which *lac* is fused has been removed, the fusions are equivalent to deletions which inactivate the gene product. The case of autogenous control is an instance when absence of the product of the gene being studied is critical to the analysis. Autogenous control refers to situations in which the product of a gene controls its own synthesis. One case of autogenous control which was revealed by genetic fusion studies is that of the *araC* gene, which codes for a positive control factor for the *ara* operon. To study this mechanism, strains diploid for the *araC* locus were constructed in which the *araC* gene introduced on an F′ episome carried different alleles (*araC*^{c}, *araC*, and *araC*$^+$). However, if the chromosomal allele had been *araC*$^+$, numerous difficulties would have arisen because of interactions between the different alleles. This difficulty was overcome and the tedious assay for *araC* protein was avoided by replacing the chromosomal *araC* gene with an *araC-lac* fusion. Assays of β-galactosidase activity in the various diploid strains showed that the *araC* gene product represses its own synthesis (20). Similarly, fusions of the *lac* genes to the *putA* gene in *Salmonella typhimurium* revealed an autoregulatory mechanism in this system (72).

Studying regulation of a gene in the absence of its product has also been important in the case of two systems in which transmembrane signaling appears to be involved. Evidence had suggested that the glucose-6-phosphate transport protein, coded for by the *uhpT* gene, was induced only by glucose-6-phosphate located outside of the cytoplasmic membrane. A complicating factor in studying this regulation was that the protein whose regulation was being studied was itself involved in transporting glucose-6-phosphate into the cytoplasm. The demonstration that glucose-6-phosphate induced the synthesis of β-galactosidase in a *uhpT-lac* fusion provided convincing evidence for the hypothesis (100). Similar studies with the regulation of genes involved in potassium transport indicated that the osmotic pressure difference between the inside and the outside of the cell was the critical factor in determining regulation (68).

Determining the Direction of Transcription of a Gene

When the direction of transcription of a bacterial gene is not known, properties of *lac* fusions to that gene provide a means of determining the orientation. Clearly, when the *lac* genes have been fused to another gene (X) so that they are under the control of the promoter and controlling elements of gene X, the *lac* operon must have the same orientation as gene X. If the fusion is on the bacterial chromosome, methods exist for determining the orientation of the *lac* operon. For instance, Freitag and Eisenstein (38) obtained fusions of the *lac* operon to a gene (*fimD*) determining the synthesis of type I fimbriae in *E. coli*. The orientation of this gene was not known. Introduction of an F-*lac* episome into such a strain and selection for integration of the F-*lac* into the chromosome via the *lac* homology resulted in an Hfr strain. The direction of transfer of chromosomal markers by the Hfr is dependent on the orientation of the chromosomal

lac region in the fusion. Determining the orientation of transfer by the Hfr obtained in this way gives the direction of transcription of the *lac* genes and, thus, of the *fim* genes.

Alternatively, deletions obtained from *lac* fusions made with the Mud *lac* phage provide a means of determining gene orientation. The Mu d phage is temperature inducible. Certain temperature-resistant mutants of Mu d lysogens which retain the *lac* fusion are due to deletions which remove the offending Mu genes and neighboring bacterial genes. Those genes which are deleted must be promoter distal to the operon. Genes on the opposite side of the fusion from Mu cannot be deleted without removing the fusion itself. Wanner et al. (119) used this approach to determine the direction of transcription of genes (*psi* genes) regulated by inorganic phosphate. A somewhat different approach based on the same principles was used to orient the *glnA* and *glnG* genes on the chromosome (71).

With fusions that have an adjacent λ prophage, the temperature-sensitive *c*I857 allele can be crossed onto the prophage (40), and temperature-resistant deletion mutants can be analyzed as described above. Conversely, a series of λ transducing phage that carry a nearby gene can be isolated. By determining whether these phages all carry *lac* or rather genes located at the opposite end of the prophage, the direction of transcription can be deduced (107).

When *lac* fusions are obtained with genes carried on plasmids, relatively straightforward restriction mapping of the fusions will reveal orientation. Or simply determining which fragment of the gene when fused to *lac* will give *lac* expression is sufficient for such a determination.

Proving that a Gene is Expressed In Vivo

In certain cases, recombinant DNA technology and sequence analysis have revealed the possible presence of a gene before it had been clearly defined genetically. In situations such as this, fusions can be used to demonstrate that the gene is actually expressed in vivo. For example, *lacZ* gene fusions were used to demonstrate that an open translational reading frame at the distal end of the *spc* operon (28) specifies a protein, PrlA (102). Similarly, an operon fusion was used to show that a sequence coding for an antisense RNA complementary to *ompF* was made in vivo (80).

Analysis of Transcription Termination Signals

There has been very little study of the nature of transcription termination signals at the ends of operons. Most of the progress in this area has come from studies of termination signals (e.g., attenuators) which are located within a group of cotranscribed genes. The reasons for this are evident—analysis of termination signals within an operon can be done by monitoring the expression of genes downstream from such signals. Ordinarily no such monitoring can be done with signals located at the ends of operons. However, by constructing strains in which the *lac* operon is placed downstream from a transcription termination region, a means of carrying out such an analysis is provided. This approach was used to study the transcription termination site at the end of the operon determining tryptophan biosynthesis (*trp*). Strain X8605 (Fig. 9) is Lac$^-$ since the *lac* operon is missing its own promoter and transcription from the *trp* promoter terminates at *trpt*. When Lac$^+$ derivatives of X8605 are selected, many are found to contain mutations affecting *trpt*. These include mutations in the structural gene for the termination factor rho and deletions of a region downstream from the last structural gene of the *trp* operon (48, 51).

These genetic studies were crucial to revealing that the termination signals were considerably downstream from where they had been previously concluded to be. The earlier suggestion of a site for transcription termination had been based on the determination of the 3′ end of the *trp* messenger RNA obtained in vivo. The deletions obtained in the selections with X8605 showed that the signal was, in fact, approximately 75 bases downstream from the site corresponding to the 3′ end of the messenger. These fusion studies necessitated a reevaluation of the data and led to the proposal that a 3′ portion of the messenger was rapidly degraded (88).

Using vectors developed by Casadaban et al. (27), it should be possible to clone termination signals from various operons in such a way as to make the expression of the *lac* operon dependent on mutations or conditions which interfere with the termination.

Defining Genes Subject to a Common Regulatory Signal

There are numerous situations when changes in the environmental conditions lead to major changes in gene expression in bacteria. Two examples which will be discussed here are the SOS response to DNA damage and the expression of numerous genes seen after starvation for inorganic phosphate. In the case of the SOS system, several genes had been found which were turned on after DNA-damaging treatments such as ultraviolet irradiation or addition of mitomycin to the growth medium. It was possible to identify these genes because mutations in them affected phenotypes known to be manifested after DNA damage. Kenyon and Walker (62) were interested in knowing whether other genes might also be regulated by the SOS response but had not been identified because they were associated with a phenotype not yet realized. To study this question, they used the Mu d*lac* phage.

Random insertions of the Mu d phage around the bacterial chromosome were obtained. These were then characterized as to whether any of them represented fusions of the *lac* operon to a bacterial gene whose expression was turned on by DNA damage. In a limited search of this sort, Kenyon and Walker found not only some of the already characterized genes, but also some new genes.

A similar approach has been taken by Wanner and McSharry (118) to find genes regulated by inorganic phosphate. In a limited search, a number of new genes subject to phosphate regulation were found. This study also revealed that certain bacterial genes respond not only to phosphate starvation but also to several other stress conditions.

A limitation to the approach described above is that essential genes will not be detected since the Mu d insertion ordinarily destroys gene activity. However, this limitation may be overcome by isolating Mu d insertions in a collection of strains which carry F's which together cover the entire bacterial chromosome.

lac FUSIONS IN EUCARYOTIC CELLS

Several laboratories have used fusions of the *lacZ* gene to yeast genes to study regulation (52, 92). Since yeasts have no transport system for lactose, the main use of the fusions is to monitor β-galactosidase activity, either by assay or on indicator plates containing XG. Nevertheless, these fusions have played a large role in the characterization of upstream activation sites essential for gene expression (50).

Lis and co-workers (70) have constructed a fusion of the *lacZ* gene to a *Drosophila* heat shock protein. In the flies containing such fusions, β-galactosidase acts as a heat shock protein. The enzyme can be readily detected either by enzyme assay of different tissues or by staining with XG. The authors suggest that this approach will greatly facilitate numerous studies on *Drosophila* gene expression and development. They propose that by using *lacZ* fusions, the tissue specificity and timing of expression of developmentally regulated genes could be assessed conveniently and with high resolution (70).

USES OF *lac* FUSIONS TO ANALYZE PROTEIN STRUCTURE AND FUNCTION

Gene fusions provide a means to covalently label target gene product sequences with a large functional fragment of β-galactosidase. This labeling allows rapid detection of the target sequences and protects them from degradation. The LacZ activity or the physical properties of β-galactosidase can be used to isolate and purify the hybrid protein and thus the target sequences, and these reagents in turn provide important tools for the biochemical analysis of the target protein itself. Moreover, gene fusions often confer unusual phenotypes that allow the design and application of unique mutant selections. A number of ingenious techniques have been developed to exploit these novel properties of gene fusions and LacZ hybrid proteins, and this relatively simple and inexpensive technology has become increasingly popular.

Identification, Purification, and Characterization of Target Gene Products

Several different methods can be used to detect and quantitate the amount of LacZ hybrid protein in a cellular extract. As described in a previous section, assays for β-galactosidase activity are simple and sensitive. In addition, the enzyme is a very large molecule and, consequently, LacZ hybrid proteins are even larger. This property can also be used for hybrid protein detection or assay. Often this can be done directly with sodium dodecyl sulfate-polyacrylamide gel even for mixtures as complex as whole-cell extracts. Alternatively, an immunological reaction with commonly available anti-β-galactosidase antibodies can be employed. Any of these methods can serve to monitor the presence of the hybrid protein during purification.

The properties of β-galactosidase also allow simple and rapid purification of LacZ hybrid proteins. Conventional methods developed for β-galactosidase itself can be used (37). Alternatively, affinity chromatography with substrate analogs or antibodies are possibilities (44). If native protein is not required, then the hybrid can be eluted directly from one or more sodium dodecyl sulfate gels, or sizing columns can be run in the presence of detergents (104, 121).

Once purified, the hybrid protein can be used as an antigen to obtain antiserum directed against the amino-terminal target gene sequences. This antiserum is then used to identify and purify the cognate wild-type target protein. Of course, this antiserum will also recognize wild-type β-galactosidase. Consequently, when the target gene is derived from *E. coli*, experiments are performed with a *lac* deletion strain. This method, originally developed by Shuman et al. (104), has seen widespread use even in commercial settings. It is quite powerful because it works even if the target gene codes for a protein that has not be identified and whose function is not understood. Indeed, Shuman first used this method to identify a previously unseen and uncharacterized transport protein, MalF. With the development of Orf vectors and methods for generating fusions in different organisms, this method can now be used to obtain antisera directed against any protein from any source.

Amino-terminal amino acid sequence analysis of the purified hybrid protein identifies the in vivo translation start site, and it aligns correctly the translational reading frame. When dealing with unknown proteins or with internal open reading frames that have been cloned into Orf vectors, such information is particularly useful. By comparing hybrid protein sequence data to those obtained with the cognate wild-type protein, it can be determined if processing occurs, and if so, the precise site can be identified. Finally, because of the stability of LacZ fusion proteins, gene fusions may be the method of choice for obtaining reliable protein sequence data in cases in which the target gene specifies an unstable protein or peptide.

Certain proteins of interest are produced only in low amounts, and this hinders purification even if the Shuman method is used. With *lac* fusions, this problem can be solved by selecting for mutants that overexpress the hybrid protein (see above). Such a mutation can then be crossed onto the cognate wild-type gene to increase its production as well. This strategy has been used by Casadaban et al. (25) to increase production of the Tn*3* transposase. A similar strategy involving Orf vectors has also been used by Guarente et al. (49) to increase production of a cloned heterologous protein.

Another problem that hampers analysis of certain interesting proteins or peptides is their inherent instability. Two clever methods involving *lacZ* gene fusions have been described that circumvent these difficulties and allow purification of reasonable amounts of the unfused, wild-type target gene product. The first, developed by Germino and Bastia (43), utilizes site-specific proteolysis. In this scheme, a fragment containing the entire coding sequence of the protein or peptide of interest is inserted into a specially constructed Orf vector. This Orf vector is similar to those described above except that its cloning site is a small open reading frame specifying an amino acid sequence of chicken collagen. Consequently, the fusion protein produced contains the collagen sequence between the target gene product and LacZ. The fusion stabilizes the target protein, and after purification, it can be cleaved from the LacZ portion by treatment with collagenase.

A second method, developed by Enquist and co-workers (120), for stabilizing rapidly degraded proteins takes advantage of the cytoplasmic aggregation that is commonly observed when certain fusion proteins or heterologous proteins are produced at very high levels in *E. coli*. The Orf vector employed for this strategy has a cloning site that is an open reading frame except for the presence of a single nonsense codon. When an appropriate construct containing the gene of interest is placed in a host that does not contain a nonsense suppressor, the protein of interest is made, but it is rapidly degraded. On the other hand, when the vector is placed in a suppressor host, a fusion protein is made, and because of the high level of expression, it is found as an aggregate in the cytoplasm. In addition, since the efficiency of suppression is only about 50%, a large amount of the protein of interest is also made. In this case, however, it is stable because it coprecipitates with the fusion protein. Thus, by simply purifying the aggregate, the wild-type protein can be recovered.

Bifunctional Hybrid Proteins

In general, the genetic event that creates a *lacZ* gene fusion destroys target gene function. However, this is not always the case. A number of gene fusions have been described that specify bifunctional hybrid proteins, i.e., the hybrid protein has both target gene product and LacZ activities. These novel constructs are quite useful because they provide an active protein covalently labeled with LacZ.

In *E. coli*, examples of proteins that can be fused to LacZ without loss of function include cytoplasmic enzymes (TrpA [78] and ThrA [93]) regulatory proteins (LacI [81], λ repressor [49], and OmpR [14; see below]), other DNA-binding proteins (replication initiator protein from plasmid R6K [41, 42]), and a protein that functions in sugar transport into the cytoplasm (MalK [36]). The LacI hybrids provided information about the relationship between subunit oligomerization and DNA-binding activity (57, 61; see the following section). The MalK hybrid was useful for identifying the wild-type protein and determining its cellular location. It was also employed to study protein-protein interactions at the cytoplasmic membrane. Results indicate that MalK is a peripheral cytoplasmic membrane protein bound to the membrane through an interaction with another maltose transport protein, MalG (103).

In yeasts, *MATα2-lacZ* and *HAP2-lacZ* gene fusions that specify bifunctional hybrid proteins have been described (55; J. Pinkham and L. Guarente, personal communication). These fusions are particularly intriguing because the target gene products are regulatory proteins. Presumably, the hybrid protein is entering the nucleus. Indeed, the MATα2-LacZ hybrid protein binds specifically to target DNA sequences in vitro (M. Hall, personal communication).

Identification of Functional Protein Domains

Even though the genetic event that creates a gene fusion may destroy the overall function of the target gene product, certain activities may remain. The LacZ portion of these hybrid proteins provides a convenient means to stabilize and purify various portions of the target protein and, thus, identify the responsible domain. For example, by determining which LacI, replication initiator, and MATα2 hybrid proteins still retain DNA-binding activity, it has been possible to identify the functional domain (41, 42, 44, 57, 61, 81; M. Hall, personal communication). Similar studies should be possible with OmpR, a regulatory protein that stimulates transcription of two different genes. In this case, gene fusions have been used to demonstrate in vivo that the carboxy terminus of the regulatory protein is required for activation of one promoter but not another (14).

Gene fusions have also been used to identify domains that participate in protein-protein interactions. Indeed, LacZ itself has been extensively studied by using fusions. By determining the amount of LacZ remaining in a variety of different hybrid proteins and correlating this with specific activity and thermal stability, important information about the residues required for dimer-dimer interaction was obtained (126). As described above, certain TrpA hybrid proteins are bifunctional. Some of those which have no TrpA activity still interact with TrpA or TrpB (M. Berman, personal communication). Closer examination of these hybrids should provide information concerning the relevant domains.

Identification of Intragenic Export Signals

Gene fusions have been widely used to study the process of protein export to the periplasm and the outer membrane of *E. coli* (11, 74, 85, 106). The rationale here is similar to that described for identifying functional domains; namely, by constructing a series of *lacZ* fusions to a gene specifying an exported protein and then determining the cellular location of the hybrid protein that is produced, insights into the nature and location of intragenic export signals can be obtained. Results of these studies were instrumental in demonstrating the importance of the amino-terminal signal sequence for protein localization (see below). They indicate that a common pathway is employed for export to either the periplasm or the outer membrane. These pathways diverge, however, since LacZ hybrids can apparently be localized to the outer membrane but not to the periplasm. In addition to the signal sequence, other intragenic information has been implicated. With the outer membrane protein LamB, at least two other signals have been identified (10). One of these may function as a sorting sequence to specifically direct the protein to the outer membrane. Consistent with this proposal is the observation that other outer membrane proteins contain a similar, homologous sequence (82). With MalE, a periplasmic protein, an additional downstream signal has also been described and mapped (5). The function of this signal is not yet clear; however, it may provide an explanation for the apparent delay in the translocation of this protein across the cytoplasmic membrane (89).

In recent years the application of *lacZ* fusions for the study of protein localization has been extended to yeasts in which target gene products are destined for secretion and a variety of different cellular locations including the nucleus, mitochondria, and the vacuole. As mentioned above, certain MATα2 and HAP2 hybrid proteins appear to be localized to the nucleus (55; J. Pinkham and L. Guarente, personal communication). Similar studies have also been reported with GAL4 (another regulatory protein) hybrids (109). Results suggest the presence of a small amino acid sequence that is common to nuclear proteins and is responsible for this routing. Although more work is required, it seems quite possible that protein localization to the nucleus is an active, receptor-mediated process.

Many mitochondrial proteins are also made initially in precursor form with an amino-terminal signal sequence. This sequence, as might be expected, differs markedly from the signal sequences described above and those present on proteins that are routed through the endoplasmic reticulum (see below). Fusions have been constructed to the nuclear gene encoding the mitochondrial matrix protein ATP2, a component of adenosine triphosphatase. These LacZ hybrids appear to be localized correctly even if relatively small sequences of target gene product are present (33). This indicates that the carboxy-terminal portion of the protein is not required for import, and, given the fact that mitochondrial protein import occurs posttranslationally, this result is intriguing.

Proteins destined for secretion and proteins destined for the vacuole are both thought to be routed through the endoplasmic reticulum. Indeed, both are synthesized initially with a typical amino-terminal signal sequence. SUC2 (invertase)-LacZ hybrid proteins that contain the signal sequence and small amounts of mature sequence appear to be exported into the endoplasmic reticulum where they remain (35). This is consistent with the proposal that proteins must assume an appropriate conformation to exit this organelle (97).

As will be discussed in the following section, it appears that sequences of β-galactosidase cannot cross the inner membrane of *E. coli*. Results obtained so far with yeasts

suggest that a similar problem does not exist. Whether or not this reflects a fundamental difference between eucaryotes and procaryotes remains to be determined.

Novel Phenotypes Conferred by Gene Fusions

One of the most useful applications of gene fusion technology for the study of protein export in *E. coli* stems from the unusual phenotypes exhibited by strains that produce an exported LacZ hybrid protein. These phenotypes fall generally into two types. The first is characteristic of strains that produce a hybrid protein that is localized incorrectly or inefficiently. Such strains exhibit a phenotype termed overproduction lethality. Because high-level synthesis of such hybrid proteins results in generalized precursor accumulation, it is though that these fusions cause a lethal jamming of the cellular export machinery when they are present in high amounts. The second characteristic phenotype is observed with strains that produce a hybrid protein that is localized to a membrane with high efficiency. In this case, the hybrid protein exhibits extremely low β-galactosidase activities. This decrease in activity is thought to reflect the inability of LacZ to oligomerize properly in this foreign environment. These unusual phenotypes are useful because they can be exploited to select directly for export-defective mutants.

Mutations identified by using selections based on the novel properties of *lacZ* fusions to genes specifying exported proteins have provided conclusive evidence for the function of the signal sequence in protein export, and they have focused attention on other intragenic regions that appear to contain additional export information. They have also been used to isolate mutants that are generally export defective. By characterizing these mutations it has been possible to identify genes that code for components of the cellular export machinery. Gene fusions were used to obtain antisera directed against certain of these components and to demonstrate that they are regulated in response to the export requirements of the cell. In addition, fusions have been instrumental in demonstrating a mechanism that couples translation and export in bacteria. A number of current reviews describing these phenotypes (6), the selections (105), and the characterization of the various mutants are available (11, 74, 85, 106).

Several unusual phenotypes have also been observed with *lacZ* fusions in yeasts. For example, synthesis of the MATα2-LacZ hybrid protein causes a pronounced growth inhibition (55). Also, certain ATP2-LacZ hybrid proteins cause a respiration-defective phenotype (33). Although the molecular mechanisms responsible for these phenotypes are not yet known, they offer obvious selections that are likely to provide a number of interesting and useful mutants.

Although other examples of unusual phenotypes associated with *lacZ* fusions are not common, we see no reason why this phenomenon cannot be generalized. For example, it should be possible to isolate fusions to certain genes that exhibit a dominant negative phenotype, and these could be exploited. Consider the *trpA-lacZ* W209 or the *recA-lacZ* fusions mentioned previously. The first fusion is $TrpA^-$; however, when expressed at high levels, it makes a Trp^+ cell Trp^- (M. Berman, personal communication). This could reflect mixed multimers of TrpA and TrpA-LacZ or it could indicate a detrimental interaction of the hybrid with TrpB. The latter fusion confers an overproduction lethality phenotype (123a). Mutant analysis in these cases would be greatly facilitated by the phenotypes associated with the *lac* portion of the fusion, and it should provide information concerning the relevant structure-function relationships. Alternatively, since fusions that specify bifunctional hybrid proteins can be obtained, it may be possible by overproduction to alter a phenotype associated with the target gene.

USES OF *lac* FUSIONS TO ANALYZE THE TARGET GENE

Besides serving as a covalent protein modification, *lac* fusions also tag the corresponding target DNA sequence, and here again, a variety of useful applications have been devised.

Physical Analysis of *Cis*-Acting Regulatory Sites

As described in a previous section, one of the chief advantages of *lac* fusion technology is the ability to isolate λ transducing phage that carry the hybrid gene. By doing this, a molecular clone of the 5′ sequences of the target gene and the *cis*-acting regulatory sites can be obtained. Using standard methods and an appropriate acceptor (107), any fusion can be subcloned onto a multicopy plasmid vector and analyzed at the level of the DNA sequence. This information alone often provides important insights into the nature and function of the regulatory signals. Moreover, by combining this biochemical characterization with thorough genetic analysis, the molecular functions of these regulatory sequences can be dissected and demonstrated. The alterations caused by various *cis*-acting regulatory mutations can be determined since they can be recombined from the transducing phage to the corresponding plasmid. On the other hand, mutations constructed on a plasmid can be recombined onto the transducing phage so that their effects on the fusion can be analyzed without fear of artifacts caused by the high copy number of the plasmid.

Cloning the Target Gene

Molecular cloning of the 5′ sequences of the target gene provide a DNA sequence, available in large quantities, that can be used as a hybridization probe to identify the cognate wild-type gene in an appropriate library. With *E. coli* genes, this application can be simplified further. Since the λ transducing phage carrying the fusion is att^-, it is possible to isolate lysogens in a *lac* deletion strain in which the prophage lies adjacent to the cognate, wild-type gene. After induction, it should be possible to identify phage that have picked up the wild-type gene by an aberrant excision event. Furthermore, since the λ *plac* Mu phage can integrate at almost any point in the chromosome, it should be possible to direct insertion in a manner similar to that used to isolate transposon insertions near a gene of interest (63, 107). After induction, transducing phage that carry the nearby gene of interest can then be isolated (16). After subcloning, the wild-type gene would then be available for DNA sequence analysis and further genetic manipulations.

Detection and Cloning of Genes That Code for Known Proteins

There are often instances in which a protein has been identified, purified, and characterized, yet its structural gene remains elusive. In these cases, *lacZ* fusions can be used to identify a portion of the target gene. This can be accomplished by using Orf vectors to clone small random DNA fragments from the organism of interest. Clones that have acquired an insert, that is, an open reading frame, are identified as those that confer a $LacZ^+$ phenotype (see above). By screening these for immunological cross-

reactivity toward the protein of interest, clones that carry a portion of the target gene can be identified. This DNA fragment can then be purified and used as a probe to identify clones which carry the entire target gene.

Although this random cloning method for detecting genes of interest seems both risky and complicated, it is quite sensitive and powerful. It has been used successfully even when the cloned sequence specifies only 10 amino acids of the target gene product (64). By employing a series of monoclonal antibodies, this analysis can be further extended, to map gene sequences that specify a particular epitope (120, 122). Thus, *lac* fusions can be used to identify any gene if an antibody directed against the gene product is available.

CONCLUSION

The analysis of biological problems has been greatly facilitated over the last 10 years by the development of powerful new genetic approaches. The use of restriction enzymes, the appearance of recombinant DNA technology, and the refinement of DNA sequencing techniques have all contributed to an impressive explosion of biological knowledge. We suggest that the use of genetic fusions is an approach to biological problems which has also been of major significance to this explosion. From relatively simple analyses such as those of gene regulation in bacteria to the study of the complex process of differentiation, the utility of genetic fusions has been demonstrated. A range of questions from protein structure-function relationships to protein secretion have all benefited from this new approach.

In this review, we have documented the wide range of uses of the genetic fusion approach. We have chosen to illustrate the principles with the *lac* fusion techniques, since we are most familiar with them and they have been the most widely used. However, other genes have been used for fusions with great success. These include the *galK* (galactokinase [73]), the *cat* (chloramphenical transacetylase [47]), the *npt* (neomycin phosphotransferase [117]), and the *phoA* (alkaline phosphatase [57a; C. Manoil and J. R. Beckwith, Proc. Natl. Acad. Sci. USA, in press]) genes. These other systems can, in some circumstances, provide advantages over the *lac* fusion system. For instance, the expression of the *E. coli galK* gene can be scored for in yeasts by using the galactose phenotype (90). Since alkaline phosphatase can pass through a membrane, but β-galactosidase cannot in *E. coli*, fusion systems with *phoA* allow the detection and manipulation of signals for export of proteins (Manoil and Beckwith, in press).

Given then, the wide range of uses to which *lac* fusions have been put and the development of new fusion systems with different uses, we imagine that the genetic fusion approach will be even more fruitful in the future.

ACKNOWLEDGMENTS

Many of our colleagues were kind enough to share their unpublished data with us and for this we are grateful. Special thanks go to Michael Berman and Robert Simons for providing Fig. 7 and 8, respectively, and to Ann McIntosh and Susan DiRenzo for their help in the preparation of this manuscript.

LITERATURE CITED

1. **Accolla, R. S., and F. Celada.** 1976. Antibody-mediated activation of a deletion mutant β-galactosidase defective in the alpha region. FEBS Lett. **67**:299–303.
2. **Askov, S., C. L. Squires, and C. Squires.** 1984. Translational coupling of the *trpB* and *trpA* genes in the *Escherichia coli* tryptophan operon. J. Bacteriol. **157**:363–367.
3. **Baker, H. V., and R. E. Wolfe.** 1983. Growth rate-dependent regulation of 6-phosphogluconate dehydrogenase level in *Escherichia coli* K-12: β-galactosidase expression in *gnd-lac* operon fusion strains. J. Bacteriol. **153**:771–781.
4. **Baker, H. V., and R. E. Wolf.** 1984. An essential site for growth rate dependent regulation within the *Escherichia coli gnd* structural gene. Proc. Natl. Acad. Sci. USA **81**:7669–7673.
5. **Bankaitis, V. A., and P. J. Bassford.** 1984. The synthesis of export-defective proteins can interfere with normal protein export in *Escherichia coli*. J. Biol. Chem. **259**:12193–12200.
6. **Beckwith, J., and T. J. Silhavy.** 1983. Genetic analysis of protein export in *Escherichia coli*. Methods Enzymol. **97**:3–11.
7. **Beckwith, J. R.** 1964. A deletion analysis of the *lac* operator region in *E. coli*. J. Mol. Biol. **8**:427–430.
8. **Beckwith, J. R., E. R. Signer, and W. Epstein.** 1966. Transposition of the *lac* region of *E. coli*. Cold Spring Harbor Symp. Quant. Biol. **31**:393–401.
9. **Bedouelle, H.** 1983. Mutations in the promoter regions of the *malEFG* and *malK-lamB* operons of *Escherichia coli* K12. J. Mol. Biol. **170**:861–882.
10. **Benson, S. A., E. Bremer, and T. J. Silhavy.** 1984. Intragenic regions required for LamB export. Proc. Natl. Acad. Sci. USA **81**:3830–3834.
11. **Benson, S. A., M. N. Hall, and T. J. Silhavy.** 1985. Genetic analysis of protein export in *Escherichia coli* K-12. Annu. Rev. Biochem. **54**:101–134.
12. **Benzer, S., and S. P. Champe.** 1962. A change from nonsense to sense in the genetic code. Proc. Natl. Acad. Sci. USA **48**:1114–1120.
13. **Berman, M., and J. Beckwith.** 1979. Use of gene fusions to isolate promoter mutants in the transfer RNA gene, *tyrT*, of *Escherichia coli*. J. Mol. Biol. **130**:305–315.
14. **Berman, M. L., and D. E. Jackson.** 1984. Selection of *lac* gene fusions in vivo: *ompR-lacZ* fusions that define a functional domain of the *ompR* gene product. J. Bacteriol. **159**:750–756.
15. **Bremer, E., E. Beck, I. Hindennach, I. Sonntag, and U. Henning.** 1980. Cloned structural gene *ompA* for an integral outer membrane protein of *Escherichia coli* K-12 localization on hybrid plasmid PTU-100 and expression of a fragment of the gene. Mol. Gen. Genet. **179**:13–20.
16. **Bremer, E., T. J. Silhavy, and G. M. Weinstock.** 1985. Transposable λ *plac* Mu bacteriophages for creating *lacZ* operon fusions and kanamycin resistance insertions in *Escherichia coli*. J. Bacteriol. **162**:1092–1099.
17. **Bremer, E., T. J. Silhavy, J. M. Weisemann, and G. M. Weinstock.** 1984. λ *plac* Mu: a transposable derivative of bacteriophage λ for creating *lacZ* protein fusions in a single step. J. Bacteriol. **158**:1084–1093.
18. **Brickman, E., T. Silhavy, P. J. Bassford, H. Shuman, and J. Beckwith.** 1979. Sites within the *lacZ* gene for formation of hybrid active β-galactosidase molecules. J. Bacteriol. **139**:13–18.
19. **Campbell, A.** 1971. Genetic structure, p. 13–44. *In* A. D. Hershey (ed.), The bacteriophage lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Casadaban, M.** 1976. Regulation of the regulatory gene for the arabinose pathway, *araC*. J. Mol. Biol. **104**:557–566.
21. **Casadaban, M., and S. N. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138**:179–207.
22. **Casadaban, M. J.** 1976. Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage λ and Mu. J. Mol. Biol. **104**:541–555.
23. **Casadaban, M. J., and J. Chou.** 1984. *In vivo* formation of gene fusions encoding hybrid β-galactosidase proteins in one step with a transposable Mu-*lac* transducing phage. Proc. Natl. Acad. Sci. USA **81**:535–539.
24. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1980. In vitro gene fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection of transla-

tional initiation signals. J. Bacteriol. **143**:971–980.
25. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1982. Overproduction of the Tn3 transposition protein and its role in DNA transposition. Cell **28**:345–354.
26. **Casadaban, M. J., and S. N. Cohen.** 1979. Lactose genes fused to exogenous promoters in one step using a Mu-*lac* bacteriophage: *in vivo* probe for transcriptional control sequences. Proc. Natl. Acad. Sci. USA **76**:4530–4533.
27. **Casadaban, M. J., A. Martinez-Arias, S. K. Shapira, and J. Chou.** 1983. β-Galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. Methods Enzymol. **100**:293–307.
28. **Cerretti, D. P., D. Dean, G. R. Davis, D. M. Bedwell, and M. Nomura.** 1983. The *spc* ribosomal protein operon of *Escherichia coli*: sequence and cotranscription of the ribosomal protein genes and a protein export gene. Nucleic Acids Res. **11**:2599–2616.
29. **Champe, S. P., and S. Benzer.** 1962. An active cistron fragment. J. Mol. Biol. **4**:288–292.
30. **Chou, J., M. J. Casadaban, P. G. Lemaux, and S. N. Cohen.** 1979. Identification and characterization of a self-regulated repressor of translocation of the Tn3 element. Proc. Natl. Acad. Sci. USA **76**:4020–4024.
31. **Crick, F. H. C., L. Barnett, S. Brenner, and R. J. Watts-Tobin.** 1961. General nature of the genetic code for proteins. Nature (London) **192**:1227–1232.
32. **Debarbouille, M., H. A. Shuman, T. J. Silhavy, and M. Schwartz.** 1978. Dominant constitutive mutations in *malT*, the positive regulator gene of the maltose regulon in *Escherichia coli*. J. Mol. Biol. **124**:359–371.
33. **Douglas, M. G., B. L. Geller, and S. D. Emr.** 1984. Intracellular targeting and import of an F1-ATPase β-subunit-β-galactosidase hybrid protein into yeast mitochondria. Proc. Natl. Acad. Sci. USA **81**:3983–3987.
34. **Ebright, R., P. Cossart, B. Gicquel-Sanzey, and J. Beckwith.** 1984. Mutations that alter the DNA-sequence of the catabolite gene activator protein (CAP) in *E. coli*. Nature (London) **311**:232–235.
35. **Emr, S. D., I. Schauer, P. Esmon, and R. Schekman.** 1984. Invertase β-galactosidase hybrid proteins fail to be transported from the endoplasmic reticulum in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4**:2347–2355.
36. **Emr, S. D., and T. J. Silhavy.** 1980. Mutations affecting localization of an *Escherichia coli* outer membrane protein, the bacteriophage λ receptor. J. Mol. Biol. **141**:63–90.
37. **Fowler, A. V., and I. Zabin.** 1983. Purification, structure, and properties of hybrid β-galactosidase proteins. J. Biol. Chem. **258**:14354–14358.
38. **Freitag, C. S., and B. I. Eisenstein.** 1983. Genetic mapping and transcriptional orientation of the *fimD* gene. J. Bacteriol. **156**:1052–1058.
39. **Froshauer, S., and J. Beckwith.** 1984. The DNA sequence of the *malF* gene and of *lacZ* fusions to it. J. Biol. Chem. **259**:10896–10903.
40. **Garrett, S., R. K. Taylor, T. J. Silhavy, and M. L. Berman.** 1985. Isolation and characterization of Δ*ompB* strains of *Escherichia coli* by a general method based on gene fusions. J. Bacteriol. **162**:840–844.
41. **Germino, J., and D. Bastia.** 1983. The replication initiator protein of plasmid R6K tagged with β-galactosidase shows sequence-specific DNA-binding. Cell **32**:131–140.
42. **Germino, J., and D. Bastia.** 1983. Interaction of the plasmid R6K-encoded replication initiator protein with its binding sites on DNA. Cell **34**:125–134.
43. **Germino, J., and D. Bastia.** 1984. Rapid purification of a cloned gene product by genetic fusion and site-specific proteolysis. Proc. Natl. Acad. Sci. USA **81**:4692–4696.
44. **Germino, J., J. G. Gray, H. Charbonneau, T. Vanaman, and D. Bastia.** 1983. Use of gene fusions and protein-protein interaction in the isolation of a biologically active regulatory protein: the replication initiator protein of plasmid R6K. Proc. Natl. Acad. Sci. USA **80**:6848–6852.
45. **Gray, M. R., H. V. Colot, L. Guarente, and M. Rosbash.** 1982. Open reading frame cloning: identification, cloning, and expression of open reading frame DNA. Proc. Natl. Acad. Sci. USA **79**:6598–6602.
46. **Gronenborn, B., and J. Messing.** 1978. Methylation of single-stranded DNA in vitro introduces new restriction endonuclease cleavage sites. Nature (London) **272**:375–377.
47. **Guan, C., B. Wanner, and H. Inouye.** 1983. Analysis of regulation of *phoB* expression using a *phoB-cat* fusion. J. Bacteriol. **156**:710–717.
48. **Guarente, L., J. Beckwith, A. M. Wu, and T. Platt.** 1979. Localization of mutations in the transcription termination region of the tryptophan operon of *E. coli*. J. Mol. Biol. **133**:189–197.
49. **Guarente, L., G. Lauer, and T. M. Roberts.** 1980. Improved methods for maximizing expression of a cloned gene: a bacterium that synthesizes rabbit β-globin. Cell **20**:543–553.
50. **Guarente, L., and T. Mason.** 1983. Heme regulates transcription of the *CYC1* gene of *S. cerevisiae* via an upstream activation site. Cell **32**:1279–1286.
51. **Guarente, L., D. H. Mitchell, and J. Beckwith.** 1977. Transcription termination at the end of the *trp* operon of *E. coli*. J. Mol. Biol. **112**:423–436.
52. **Guarente, L., and M. Ptashne.** 1981. Fusion of *Escherichia coli lacZ* to the cytochrome c gene of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **78**:2199–2203.
53. **Gutierrez, C., and O. Raibaud.** 1984. Point mutations that reduce the expression of *malPQ*, a positively controlled operon of *Escherichia coli*. J. Mol. Biol. **177**:69–86.
54. **Hall, M. N., M. Gabay, M. Debarbouille, and M. Schwartz.** 1982. A role for secondary structure in the control of translation initiation. Nature (London) **295**:616–618.
55. **Hall, M. N., L. Hereford, and I. Herskowitz.** 1984. Targeting of *E. coli* β-galactosidase to the nucleus in yeast. Cell **36**:1057–1065.
56. **Hall, M. N., M. Schwartz, and T. J. Silhavy.** 1982. Sequence information with the *lamB* gene is required for proper routing of the bacteriophage λ receptor protein to the outer membrane of *Escherichia coli* K-12. J. Mol. Biol. **156**:93–112.
57. **Heidecker, G., and B. Muller-Hill.** 1977. Synthetic multifunctional proteins. Mol. Gen. Genet. **155**:301–307.
57a. **Hoffman, C., and A. Wright.** 1985. Fusions of secreted proteins to alkaline phosphatase: a new approach for studying protein secretion. Proc. Natl. Acad. Sci. USA **82**:5107–5111.
58. **Hopkins, J.** 1974. A new class of promoter mutation in the lactose operon of *Escherichia coli*. J. Mol. Biol. **87**:715–724.
59. **Huisman, O., and R. D'Ari.** 1983. Effects of suppressors of SOS-mediated filamentation on *sfiA* operon expression in *Escherichia coli*. J. Bacteriol. **153**:169–175.
60. **Jacob, F., A. Ullman, and J. Monod.** 1965. Deletions fusionmant l'operon lactose et un operon purine chez *Escherichia coli*. J. Mol. Biol. **42**:511–520.
61. **Kania, J., and B. Muller-Hill.** 1977. Construction, isolation and implications of repressor-galactosidase:β-galactosidase hybrid molecules. Eur. J. Biochem. **79**:381–386.
62. **Kenyon, C. J., and G. C. Walker.** 1980. DNA-damaging agents stimulate gene expression at specific loci in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **77**:2819–2823.
63. **Kleckner, N., J. Roth, and D. Botstein.** 1977. Genetic engineering *in vivo* using translocatable drug-resistance elements: new methods in bacterial genetics. J. Mol. Biol. **116**:125–159.
64. **Koenen, M., U. Ruther, and B. Muller-Hill.** 1982. Immunoenzymatic detection of expressed gene fragments cloned in the *lacZ* gene of *E. coli*. EMBO J. **1**:509–512.
65. **Komeda, Y., and T. Iino.** 1979. Regulation of expression of the flagellin gene (*hag*) in *Escherichia coli* K-12: analysis of *hag-lac* gene fusions. J. Bacteriol. **139**:721–729.
66. **Kroos, L., and D. Kaiser.** 1984. Construction of Tn *lac*, a transposon that fuses *lacZ* expression to exogenous promoters, and its introduction into *Myxococcus xanthus*. Proc. Natl. Acad. Sci. USA **81**:5816–5820.
67. **Kuritzkes, D. R., X. Zhang, and E. C. C. Lin.** 1984. Use of Φ*glp-lac* in studies of respiratory regulation of the *Escherichia coli* anaerobic *sn*-glycerol-3-phosphate dehydrogenase genes

(*glpAB*). J. Bacteriol. **157**:591–598.
68. **Laimins, L. A., D. B. Rhoads, and W. Epstein.** 1981. Osmotic control of *kdp* expression in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **78**:464–468.
69. **Leather, T. D., J. Noti, and H. E. Umbarger.** 1979. Physical characterization *ilv-lac* fusions. J. Bacteriol. **140**:251–260.
70. **Lis, J. T., J. A. Simon, and C. A. Sutton.** 1983. New heat-shock puffs and β-galactosidase activity resulting from transformation of *Drosophila* with an *hsp70-lacZ* hybrid gene. Cell **35**:403–410.
71. **MacNeil, D.** 1981. General method, using Mu-Mud-1 dilysogens, to determine the direction of transcription of and generate deletions in the *glnA* region of *Escherichia coli*. J. Bacteriol. **146**:260–268.
72. **Maloy, S. R., and J. R. Roth.** 1983. Regulation of proline utilization in *Salmonella typhimurium*: characterization of *put*:Mu d(Ap *lac*) operon fusions. J. Bacteriol. **154**:561–568.
73. **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, p. 384–415. *In* J. S. Chirikjian and T. S. Papas (ed.), Gene amplification and analysis. Elsevier Science Publishing, Inc., New York.
74. **Michaelis, S., and J. Beckwith.** 1982. Mechanism of incorporation of cell envelope proteins in *Escherichia coli*. Annu. Rev. Microbiol. **36**:435–464.
75. **Michaelis, S., L. Guarente, and J. Beckwith.** 1983. In vitro construction and characterization of *phoA-lacZ* gene fusions in *Escherichia coli*. J. Bacteriol. **154**:356–365.
76. **Miller, J. H., W. S. Reznikoff, A. E. Silverstone, K. Ippen, E. R. Signer, and J. R. Beckwith.** 1970. Fusions of the *lac* and *trp* regions of the *E. coli* chromosome. J. Bacteriol. **104**:1273–1279.
77. **Miller, V. L., and J. J. Mekalanos.** 1984. Synthesis of cholera toxin is positively regulated at the transcriptional level by *toxR* Proc. Natl. Acad. Sci. USA **81**:3471–3475.
78. **Mitchell, D., W. Reznikoff, and J. Beckwith.** 1975. Genetic fusions defining *trp* and *lac* regulatory elements. J. Mol. Biol. **93**:331–350.
79. **Miura, A., H. Krueger, S. Itoh, H. A. deBoer, and M. Normura.** 1981. Growth rate dependent regulation of ribosome synthesis in *E. coli*: expression of the *lacZ* and *galK* genes fused to ribosomal promoters. Cell **25**:773–782.
80. **Mizuno, T., M. Chou, and M. Inouye.** 1984. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). Proc. Natl. Acad. Sci. USA **81**:1966–1970.
81. **Muller-Hill, B., and J. Kania.** 1974. *lac* repressor can be fused to β-galactosidase. Nature (London) **249**:561–563.
82. **Nikaido, H., and H. C. P. Wu.** 1984. Amino acid sequence homology among the major outer membrane proteins of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **81**:1048–1052.
83. **O'Connor, M. B., and M. H. Malamy.** 1983. A new insertion sequence, IS121, is found on the Mu dII(Ap *lac*) bacteriophage and the *Escherichia coli* chromosome. J. Bacteriol. **156**:669–679.
84. **O'Day, K. J., D. W. Schultz, and M. M. Howe.** 1978. Search for integration-deficient mutants of bacteriophage Mu, p. 48–51. *In* D. Schlessinger (ed.), Microbiology—1978. American Society for Microbiology, Washington, D.C.
85. **Oliver, D.** 1985. Protein secretion in *Escherichia coli*. Annu. Rev. Microbiol. **39**:615–648.
86. **Oliver, D., and J. Beckwith.** 1981. *E. coli* mutant pleiotropically defective in the secretion of *E. coli* proteins. Cell **25**:765–772.
87. **Oppenheim, D. S., and C. Yanofsky.** 1980. Translational coupling during expression of the tryptophan operon in *Escherichia coli*. Genetics **95**:785–795.
88. **Platt, T., and D. G. Bear.** 1983. Role of RNA-polymerase, rho factor and ribosomes in transcription termination, p. 123–161. *In* J. Beckwith, J. Davies, and J. Gallant (ed.), Gene function in prokaryotes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
89. **Randall, L. L., and S. J. S. Hardy.** 1984. Export of protein in bacteria. Microbiol. Rev. **48**:290–298.
90. **Rymond, B. R., R. S. Zitomer, D. Schumperli, and M. Rosenberg.** 1983. The construction and expression in yeast of *CYC1-galK* fusion genes. Gene **25**:249–262.
91. **Reznikoff, W., J. H. Miller, J. G. Scaife, and J. R. Beckwith.** 1969. A mechanism for repressor action. J. Mol. Biol. **43**:201–213.
92. **Rose, M., M. Casadaban, and D. Botstein.** 1981. Yeast genes fused to β-galactosidase in *Escherichia coli* can be expressed normally in yeast. Proc. Natl. Acad. Sci. USA **78**:2460–2464.
93. **Saint-Girons, I.** 1978. New regulatory mutations affecting the expression of the threonine operon in *Escherichia coli* K-12. Mol. Gen. Genet. **162**:95–100.
94. **Sallis, L., R. J. Hasebeck, and W. D. Nunn.** 1983. Regulation of fatty acid transport in *Escherichia coli*: analysis by operon fusion. J. Bacteriol. **155**:1450–1454.
95. **Sarthy, A., A. Fowler, I. Zabin, and J. Beckwith.** 1979. The use of gene fusions to determine a partial signal sequence of alkaline phosphatase. J. Bacteriol. **139**:932–939.
96. **Sarthy, A., S. Michaelis, and J. Beckwith.** 1981. Deletion map of the *Escherichia coli* structural gene for alkaline phosphatase. J. Bacteriol. **145**:288–292.
97. **Schauer, I., S. Emr, C. Gross, and R. Schekman.** 1985. Invertase signal and mature sequence substitutions that delay intercompartmental transport of active enzyme. J. Cell Biol. **100**:1664–1675.
98. **Schnaitman, C. A., and G. A. McDonald.** 1984. Regulation of outer membrane protein synthesis in *Escherichia coli* K-12: deletion of *ompC* affects expression of the OmpF protein. J. Bacteriol. **159**:555–563.
99. **Shapiro, J. A., and S. L. Adhya.** 1969. The galactose operon of *E. coli* K-12. II. A deletion analysis of operon structure and polarity. Genetics **62**:249–264.
100. **Shattuck-Eidens, D. M., and R. J. Kadner.** 1981. Exogenous induction of the *Escherichia coli* hexose phosphate transport system defined by *uhp-lac* operon fusions. J. Bacteriol. **148**:203–209.
101. **Shimada, K., R. A. Weisberg, and M. E. Gottesman.** 1972. Prophage λ at unusual chromosomal locations. II. Mutations induced by phage λ in *E. coli* K-12. J. Mol. Biol. **80**:297–314.
102. **Shultz, J., T. J. Silhavy, M. L. Berman, N. Fiil, and S. D. Emr.** 1982. A previously unidentified gene in the *spc* operon of *Escherichia coli* K-12 specifies a component of the protein export machinery. Cell **31**:227–235.
103. **Shuman, H. A., and T. J. Silhavy.** 1981. Identification of the *malK* gene product: a peripheral membrane component of the *Escherichia coli* maltose transport system. J. Biol. Chem. **256**:560–562.
104. **Shuman, H. A., T. J. Silhavy, and J. R. Beckwith.** 1980. Labeling of proteins with β-galactosidase by gene fusion: identification of a cytoplasmic membrane component of the *Escherichia coli* maltose transport system. J. Biol. Chem. **255**:168–174.
105. **Silhavy, T. J., and J. Beckwith.** 1983. Isolation and characterization of mutants of *Escherichia coli* K-12 affected in protein localization. Methods Enzymol. **97**:11–40.
106. **Silhavy, T. J., S. A. Benson, and S. D. Emr.** 1983. Mechanisms of protein localization. Microbiol. Rev. **47**:313–344.
107. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
108. **Silhavy, T. J., E. Brickman, P. J. Bassford, M. J. Casadaban, H. A. Shuman, V. Schwartz, L. Guarente, M. Schwartz, and J. Beckwith.** 1979. Structure of the *malB* region in *Escherichia coli* K-12. II. Genetic map of the *malEFG* operon. Mol. Gen. Genet. **174**:249–259.
109. **Silver, P. A., L. P. Keegan, and M. Ptashne.** 1984. Amino terminus of the yeast *GAL4* gene product is sufficient for nuclear localization. Proc. Natl. Acad. Sci. USA **81**:5951–5955.
110. **Simons, R. W., and N. Kleckner.** 1983. Translational control of IS*10* transposition. Cell **34**:683–691.
111. **Smith, T. F., and J. R. Sadler.** 1971. The nature of lactose

operator constitutive mutations. J. Mol. Biol. **59**:273–305.
112. **Sodergren, E. J., J. Davidson, R. K. Taylor, and T. J. Silhavy.** 1985. Selection for mutants altered in the expression or export of the outer membrane porin OmpF. J. Bacteriol. **162**:1047–1053.
113. **Taylor, R. K., S. Garrett, E. Sodergren, and T. J. Silhavy.** 1985. Mutations that define the promoter of *ompF*, a gene specifying a major outer membrane porin protein. J. Bacteriol. **162**:1054–1060.
114. **Taylor, R. K., M. N. Hall, L. Enquist, and T. J. Silhavy.** 1981. Identification of OmpR: a positive regulatory protein controlling expression of the major outer membrane matrix porin proteins of *Escherichia coli* K-12. J. Bacteriol. **147**:255–258.
115. **Tiedeman, A. A., and J. M. Smith.** 1984. Isolation and characterization of regulatory mutations affecting the expression of the *guaBA* operon in *Escherichia coli* K-12. Mol. Gen. Genet. **195**:77–82.
116. **Toussaint, A., and A. Risibois.** 1983. Phage Mu: transposition as a life style, p. 105–158. *In* J. A. Shapiro (ed.), Mobile genetic elements. Academic Press, Inc., New York.
117. **Van den Broeck, G., M. P. Timko, A. P. Kausch, A. R. Cashmore, M. Van Montagu, and L. Herrera-Estrella.** 1985. Targeting of a foreign protein to chloroplasts by fusion to the transit peptide from the small subunit of ribulose 1,5-bisphosphate carboxylase. Nature (London) **313**:358–363.
118. **Wanner, B. L., and R. McSharry.** 1982. Phosphate controlled gene expression in *Escherichia coli* K-12 using Mud-directed *lacZ* fusions. J. Mol. Biol. **158**:347–363.
119. **Wanner, B. L., S. Wieder, and R. McSharry.** 1981. Use of bacteriophage transposon Mu d1 to determine the orientation for three *proC*-linked phosphatate starvation-inducible (*psi*) genes in *Escherichia coli* K-12. J. Bacteriol. **146**:93–101.
120. **Watson, R. J., J. H. Weis, J. S. Salstrom, and L. W. Enquist.** 1984. Bacterial synthesis of herpes simplex virus types I and II glycoprotein D antigens. J. Invest. Dermatol. **88**:102–111.
121. **Weinstock, G. M., C. ap Rhys, M. L. Berman, B. Hampar, D. Jackson, T. J. Silhavy, J. Weisemann, and M. Zweig.** 1983. Open reading frame expression vectors: a general method for antigen production in *Escherichia coli* using protein fusions to β-galactosidase. Proc. Natl. Acad. Sci. USA **80**:4432–4436.
122. **Weis, J. H., L. W. Enquist, J. S. Salstrom, and R. J. Watson.** 1983. An immunologically active chimaeric protein containing herpes simplex type 1 glycoprotein D. Nature (London) **302**:72–74.
123. **Weisemann, J. M., C. Funk, and G. M. Weinstock.** 1984. Measurement of in vivo expression of the *recA* gene of *Escherichia coli* by using *lacZ* gene fusions. J. Bacteriol. **160**:112–121.
123a. **Weisemann, J. M., and G. M. Weinstock.** 1985. Direct selection of mutations reducing transcription or translation of the *recA* gene of *Escherichia coli* with a *recA-lacZ* protein fusion. J. Bacteriol. **163**:748–755.
124. **Youngman, P., P. Zuber, J. B. Perkins, K. Sandman, M. Igo, and R. Losick.** 1985. New ways to study developmental genes from spore-forming bacteria. Science **228**:285–291.
125. **Zabeau, M., and K. K. Stanley.** 1982. Enhanced expression of cro-β-galactosidase fusion proteins under the control of the p_R promoter of bacteriophage λ. EMBO J. **1**:1217–1224.
126. **Zabin, I.** 1982. β-Galactosidase alpha-complementation: a model of protein-protein interaction. Mol. Cell. Biochem. **49**:87–96.
127. **Zabin, I., A. V. Fowler, and J. R. Beckwith.** 1978. Position of the mutation in β-galactosidase ochre mutant U118. J. Bacteriol. **133**:437–438.