JOURNAL OF BACTERIOLOGY, Aug. 1980, p. 971-980
0021-9193/80/08-0971/10$02.00/0

Vol. 143, No. 2

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 and Cloning of Translational Initiation Signals

MALCOLM J. CASADABAN,*† JOANY CHOU,† AND STANLEY N. COHEN

Departments of Genetics and Medicine, Stanford University, Stanford, California 94305

We report the construction and use of a series of plasmid vectors suitable for the detection and cloning of translational control signals and 5′ coding sequences of exogenously derived genes. In these plasmids, the first eight codons of the amino-terminal end of the lactose operon β-galactosidase gene, *lacZ*, were removed, and unique *Bam*HI, *Eco*RI, and *Sma*I (*Xma*I) endonuclease cleavage sites were incorporated adjacent to the eighth codon of *lacZ*. Introduction of deoxyribonucleic acid fragments containing appropriate regulatory signals and 5′ coding sequences into such *lac* fusion plasmids led to the production of hybrid proteins consisting of the carboxyl-terminal segment of a β-galactosidase remnant plus a peptide fragment that contained the amino-terminal amino acids encoded by the exogenous deoxyribonucleic acid sequence. These hybrid peptides retained β-galactosidase enzymatic activity and yielded a Lac$^+$ phenotype. Such hybrid proteins are useful for purifying peptide sequences encoded by exogenous deoxyribonucleic acid fragments and for studies relating the structure and function of specific peptide segments.

Previously, we have reported the construction and use of plasmid vectors suitable for the cloning in *Escherichia coli* of endonuclease-generated fragments carrying transcriptional promoter or termination signals derived from either *E. coli* or foreign DNA (11, 14-17). In these plasmids, *lac* gene expression is brought under the control of transcriptional signals within the cloned DNA fragment; the Lac$^+$ phenotype of the resulting clones and the level of β-galactosidase synthesis can then be used to indicate and measure function of the exogenously derived promoters.

Since the amino-terminal end of the β-galactosidase protein is not essential for enzymatic activity (34), "translational" *lac* fusions that replace the corresponding 5′-coding segment of the *lacZ* gene with other DNA sequences, but which still have functional β-galactosidase activity, can also be constructed (4, 9, 12, 22, 34). At least the first 27 amino acid codons of *lacZ* can be removed, and the remainder of the gene can be joined to a DNA segment that encodes the regulatory signals and amino-terminal region of another gene, yielding a hybrid protein that retains the enzymatically active β-galactosidase remnant and has the same specific activity as the intact enzyme (40, 47). The length of the amino-terminal segment fused to *lacZ* can vary from as few as six amino acid codons (14, 15, 17) to a segment almost as large as an entire gene (34).

Previous methods for constructing translational *lac* gene fusions have required a series of in vivo genetic steps such as rearrangement of the genes to be fused and selection for the rare deletions that produce the fusions. Usually, only fusions that encode enough β-galactosidase activity for growth on lactose can be selected. Here we report a method for directly joining the amino-terminal segments of other genes to *lacZ* in vitro, using DNA cloning techniques, and the construction of plasmid vectors suitable for this purpose. The procedure described can be used to obtain fusions at precise locations independently of the amount of β-galactosidase produced. Since the gene fusions are made in vitro at restriction endonuclease-generated cleavage sites, the method is not limited to genes of *E. coli*; it potentially enables the joining of regulatory regions and 5′ coding sequences for a variety of eucaryotic proteins to a sequence that specifies an enzymatically active β-galactosidase gene product.

MATERIALS AND METHODS

Procedures for bacterial growth (32); isolation of plasmid DNA (26, 46); construction, cloning, and re-

† Present address: Department of Biophysics and Theoretical Biology, University of Chicago, Chicago, IL 60637.

striction endonuclease cleavage and gel analysis of recombinant plasmids (13, 18, 29, 41); transformation of *E. coli* with plasmid DNA (19); and β-galactosidase assay and purification (32) have been described previously. DNA sequence analysis employing ^{32}P-labeling of 5′ ends was carried out by the method of Maxam and Gilbert (30).

Strain MC1000 or its K-12 restriction-minus derivative MC1061 (11), both deleted for the lactose operon Δ(*lacIPOZY*)*X*74, were used for plasmid transformation and DNA cloning. Only the *lacZ* (β-galactosidase) and *Y* (lactose permease) genes need be introduced and expressed for the utilization of lactose (Lac$^+$ phenotype). The Lac phenotype was monitored on lactose-MacConkey indicator agar plates or, for very low levels of expression, on M63-glucose plates containing the chromogenic β-galactosidase substrate 5-bromo-4-chloro-3-indolyl-β-D-galactoside (6, 32). Antibiotic concentrations used were: ampicillin (Ap), 25 μg/ml; kanamycin (Km), 20 μg/ml.

Details of plasmid constructions. (i) Plasmid pMC874. Kanamycin resistance plasmid pMC275 (see Fig. 1), derived from plasmids pACYC177 (13) and pMC81 (11), was digested with *Bam*HI endonuclease, and the recessed 3′ ends were extended with reverse transcriptase (24) to yield blunt-ended molecules (1, 13). This DNA was divided into five portions (10 μg each), treated with *Hae*III endonuclease that had been serially diluted so as to yield partial digests, and electrophoresed on 0.7% agarose gels. DNA fragments that had been cleaved only once at a *Hae*III site near the *Bam*HI site were isolated as bands migrating slightly faster in gels than linear pMC275 DNA. This DNA was electrophoretically eluted, ligated, and used to transform *lac* deletion strain MC1000 (11) to Km^r, using lactose-MacConkey indicator agar containing kanamycin.

Of 74 transformed colonies, 25 were Lac$^-$ and presumably contained deletions extending into *lacZ*. Two of these could recombine (32) with all of the *lacZ* mutations tested that map after the eighth codon of *lacZ*, including the *lacZ* U118 mutation in the 17th codon (47) and the *lacZ* M15 deletion, which removes codons 11 through 41 (28). Plasmid DNA from these two was isolated and examined by gel analysis after endonuclease digestion. These plasmids were found to be slightly smaller than the original pMC275, to have a unique *Bam*HI site, to have lost several *Hae*III fragments, and to have a *Bam*HI-*Hae*III fragment which matched the size of the 102-nucleotide *lacZ* *Hae*III fragment adjacent on the downstream side to the *Hae*III site in the eighth codon of *lacZ*. One of these plasmids was designated pMC874.

(ii) Plasmid pMC871. Plasmid pMC871, which is an analog of pMC874 suitable for detection of transcriptional fusions, was constructed by homologous recombination (see Fig. 4) between plasmids pMC874 and pMC489 (11). To select for recombination between the amino-terminal portion of *lacZ* on pMC489 and the carboxyl-terminal portion of *lacZ* on pMC874, a fragment containing a promoter for *lacZ* expression was first inserted into the *Bam*HI site of pMC489. To provide this fragment, we used *Bam*HI-cleaved plasmid pMB9 DNA (7) which includes the promoter for a tetracycline resistance gene upstream from its solitary *Bam*HI cleavage site. The resulting pMC489-pMB9 hybrid plasmid, designated pMC511, was used to transform MC1000 (Δ*lac*) cells containing pMC874. Colicin E1-immune (encoded by the pMB9 segment) transformants were selected on lactose-MacConkey indicator plates. After prolonged (2 to 4 days) incubation, Lac$^+$ sectored colonies that had undergone recombination between the *lacZ* segments of the two plasmids appeared. Plasmid DNA from one of these recombinants was digested with *Bam*HI endonuclease to separate the recombined replicons and also to remove the pMB9-derived DNA segment. After ligation of the resulting fragments and transformation of MC1000, plasmid pMC871 was isolated from a Km^r colicin-sensitive transformant that showed a low level of *lacZY* expression (light-red colonies) on lactose-MacConkey plates.

(iii) Plasmids pMC1396 and pMC1403. Plasmids pMC1396 and pMC1403 are pMC874 and pBR322 (45) derivatives, respectively, containing a chemically synthesized *Eco*RI-*Sma*I (*Xma*I)-*Bam*HI sequence (3) adjacent to the *lac* segment (see Fig. 6). The synthetic sequence inserted into the *Bam*HI site of pBR322 was obtained from L. Lau and R. Wu as part of plasmid pLL10 (see Fig. 6 top); insertion of this sequence into pBR322 results in ampicillin resistance and tetracycline sensitivity. The additional *Eco*RI, *Bam*HI, and *Sma*I site on pLL10 were eliminated by removing its small *Eco*RI fragment, yielding pMC1364, which has unique adjacent *Eco*RI, *Sma*I, and *Bam*HI sites. The *Bam*HI-*Sal*I *lac* fragment of pMC874 was then introduced into pMC1364 to form pMC1396 (see Fig. 6 bottom). pMC1396 was selected as an Ap^r Km^s light-blue colony on 5-bromo-4-chloro-3-indolyl-β-D-galactoside plates after transformation of strain MC1000 (Δ*lac*) with pMC874 and pMC1364 plasmid DNAs which had been digested with *Bam*HI endonuclease plus *Sal*I endonuclease and religated. Although pMC1396 contains *Eco*RI and *Sma*I sites adjacent to the *Bam*HI site next to *lac*, the *Eco*RI site is not unique, since there is also an *Eco*RI site within the *lacZ* gene. In the formation of the pMC1403 plasmid, the *Eco*RI site was eliminated by using pMC1388, a derivative of pMC874, that contains a mutated DNA sequence in the *Eco*RI site of *lacZ*; this mutation did not affect the β-galactosidase protein. Introduction of the mutated *Eco*RI site into pMC1388 was accomplished by homologous recombination of pMC874 with λRS205, a *lac* transducing phage containing a mutated *Eco*RI site (obtained from K. Bertrand and W. Reznikoff). The recombinant was obtained by growing pMC874 in a lysogen of λRS205, isolating the plasmid DNA, and digesting it with *Eco*RI endonuclease to cleave the pMC874 plasmid DNA that had not undergone recombination. This DNA was used to transform strain MC1000 to Km^r; all of the transformants were pooled, and the total DNA was isolated and redigested with *Eco*RI. pMC1388 was obtained from this DNA by selecting a light-blue Km^r transformed colony on 5-bromo-4-chloro-3-indolyl-β-D-galactoside plates containing kanamycin.

RESULTS

Experimental strategy for construction of *lacZ* translation fusion vectors. Construc-

tion of the cloning vectors for joining amino-terminal segments of exogenous genes to *lacZ* is described in detail above. The strategy we used involved the introduction of a unique restriction endonuclease cleavage site within the early non-essential part of *lacZ*. Since there is no unique cleavage site for a known restriction endonuclease within the first 27 codons of *lacZ*, we converted a site cleaved by the *Hae*III endonuclease, which cleaves in *lacZ* in the eighth codon as well as at many other sites, into a site for another endonuclease (*Bam*HI) that did not previously cleave the *lacZ* gene.

As detailed above, we started with a plasmid containing all of *lacZ* (pMC275, Fig. 1a). The *Bam*HI site was brought into continuity with the *Hae*III site in the eighth codon of *lacZ* by deletion of the DNA between the two sites (Fig. 1b); since this *Hae*III site cannot be cleaved uniquely, deletions were made between the

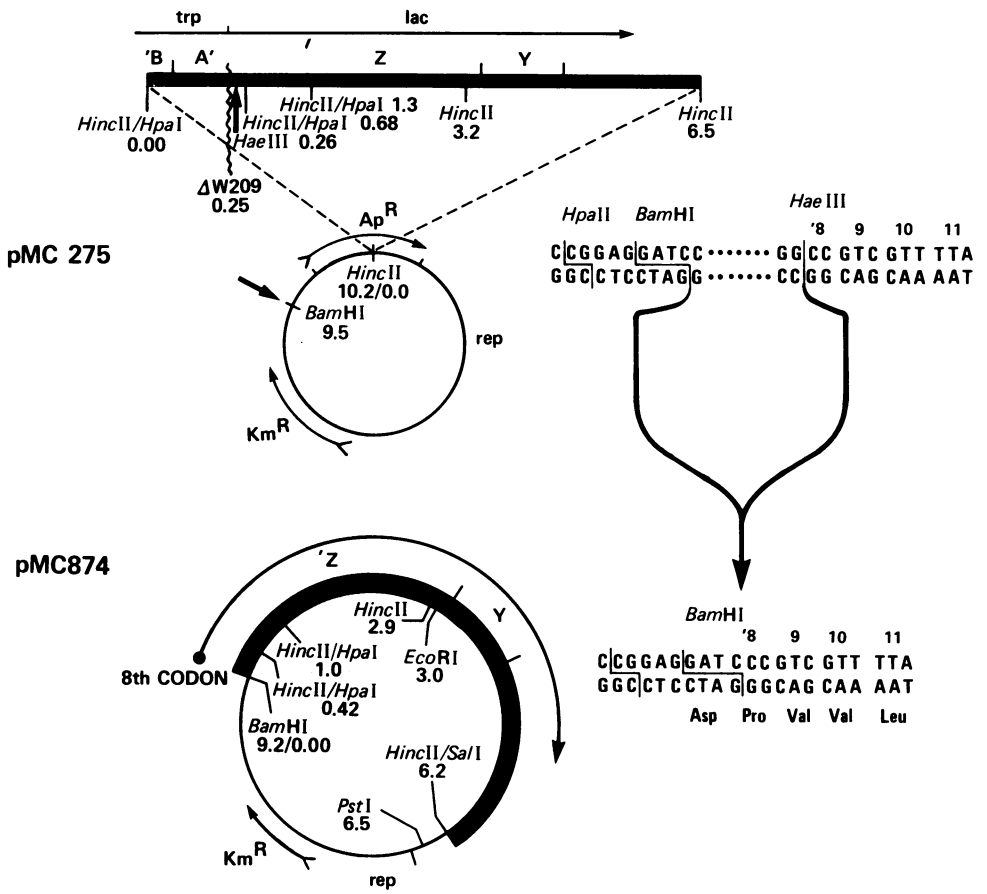

FIG. 1. *Construction of the pMC874 lac fusion vector. pMC874 was constructed from lacZ⁺Y⁺ plasmid pMC275 by deletion of DNA between a unique BamHI site and a HaeIII site located in the eighth codon of lacZ. pMC275 contains a partial HincII digestion fragment (top heavy line) of the W209 trp-lac fused operon segment (33) on pMC81 (11) originally derived from a φ80 transducing phage (8). The positions of trp operon gene segments 'B and A' and lac genes Z and Y are shown. The lac promoter is missing from this fragment and is replaced by an internal segment of the trp operon. This HincII-generated lac fragment was inserted into the unique HincII cleavage site in the bla gene encoding Apʳ of vector pACYC177 (13, 45), such that the lacZ and Y genes were expressed from the bla gene promoter. The BamHI site and the Apʳ determination of pACYC177 were derived from the Tn3 element (13, 14). A SalI site was formed after lacY at the HincII joint. pMC874 was formed by a deletion, isolated as described in the text, between the arrows on pMC275. The heavy line represents the segment derived from the lac region of pMC81. The nucleotide sequences at the sites of the deletion and of the resulting junction are given on the right along with several of the adjacent lacZ codons. The sequence of the filled in BamHI-HaeIII joint was directly confirmed by DNA sequence analysis (30), using 5' end labeling of the nearby HpaII site shown in the figure. The locations of rep (a segment essential for plasmid DNA replication) and of the Apʳ and Kmʳ determinant genes are shown. Arrows on the plasmids indicate the direction of transcription. Kilobase coordinates for relevant sites are shown.*

*Bam*HI site and different *Hae*III sites, and a deletion of the desired length was obtained by eluting the cleaved DNA from the appropriate region of agarose gels (see above). The *Bam*HI site was reformed at the site of the deletion by filling in the 5′ protruding *Bam*HI ends, using reverse transcriptase, and ligating the resulting blunt ends to the blunt ends generated by the *Hae*III endonuclease, yielding the pMC874 plasmid. Nucleotide sequence analysis at the deletion site of this plasmid was done to directly confirm the location of the deletion termini (Fig. 1). This sequence included a *Bam*HI cleavage site; the nucleotides to the right of this site match the *lacZ* gene sequence from the position of the eighth codon, and the nucleotides to the left of the *Bam*HI site match internal sequences of the Tn*3* repressor gene (14), which are present on the parental pACYC177 (13) plasmid used to construct pMC874.

Use of plasmid pMC874 for production of hybrid β-galactosidase. *E. coli* chromosomal DNA was cleaved with the *Bam*HI endonuclease and ligated with *Bam*HI cleaved pMC874 plasmid DNA (Table 1). The composite molecules were then introduced by transformation into *E. coli* strain MC1000, which carries a *lac* deletion, yielding 10^5 Km^r transformants per μg of DNA. In a control experiment, the ends of *Bam*HI-cleaved pMC874 DNA were ligated in the absence of *E. coli* chromosomal DNA, and the plasmid was introduced by transformation into MC1000. Only Lac^- transformants were obtained with the pMC874 plasmid (Table 1), consistent with the absence of a transcriptional promoter and translational initiation site. In contrast, approximately 1% of the Km^r colonies obtained from the chromosomal ligated DNA were Lac^+ and formed red colonies of various intensities on lactose-MacConkey agar containing kanamycin. Two clones that yielded dark-red colonies were selected for further study.

TABLE 1. *Use of pMC874 to form hybrid β-galactosidase proteins*[a]

E. coli chromosomal DNA	*Bam*HI digestion	Ligation	Km^r transformants/ μg of DNA	Frequency of Lac^+/ Km^r transformants
−	−	−	10^6	$<10^{-6}$
−	+	−	10^2	$<10^{-2}$
−	+	+	10^5	$<10^{-5}$
+	+	+	10^5	10^{-3}

[a] pMC874 DNA (40 μg/ml) was mixed with *E. coli* chromosomal DNA (40 μg/ml) and digested with *Bam*HI or ligated with T4 ligase or both, as listed in the table. Km^r transformants of strain MC1000 (Δ*lac*) were selected on lactose-MacConkey indicator plates containing kanamycin.

Plasmid DNA from the two clones was treated with *Bam*HI endonuclease and examined by gel electrophoresis for the presence of an adventitious DNA fragment (Fig. 2). DNA fragments 0.15 and 5 kilobases in size were found to be inserted into plasmid pMC874.

Acrylamide gel electrophoresis of proteins in crude extracts of bacteria grown from a clone carrying only the pMC874 vector showed no detectable peptide band (Fig. 3, lane f) at the location occupied by the 135,000-molecular-weight (1,021 amino acids [20]) β-galactosidase protein as shown in isopropyl-β-D-thiogalactopyranoside-induced control cells containing an intact wild-type *lacZ* gene (Fig. 3, lanes a and b). The two clones that showed a Lac^+ phenotype and contained plasmids that included an extra *Bam*HI-generated fragment also failed to show a peptide band in the normal position for β-galactosidase. However, both clones produced new proteins that were absent in extracts of cells

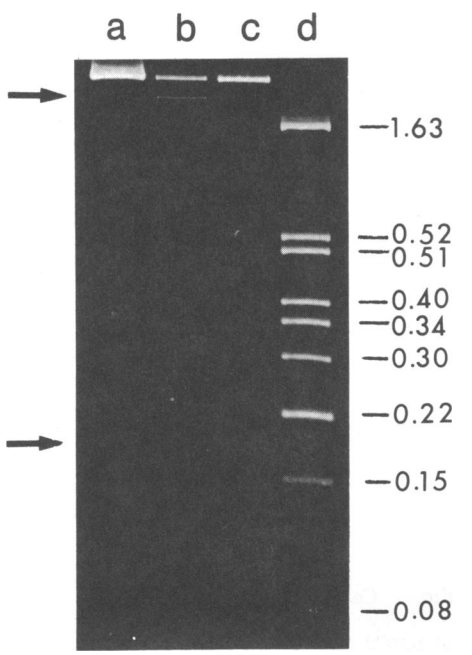

FIG. 2. *Gel analysis of translational fusion plasmids containing exogenous DNA fragments that yield Lac^+ phenotypes. Analysis of endonuclease digests of plasmid DNA was carried out on 4% polyacrylamide gels (29, 41). (a) BamHI-cleaved pMC874 vector DNA. (b and c) BamHI digests of plasmids from Lac-expressing clones, showing inserted E. coli chromosomal BamHI-generated DNA fragments (arrows). (d) Molecular size standards: HinfI-cleaved pBR322 DNA (45). The size of each fragment is indicated in kilobases. Agarose (0.7%) gels were also run to better determine the size of the large fragment (data not presented).*

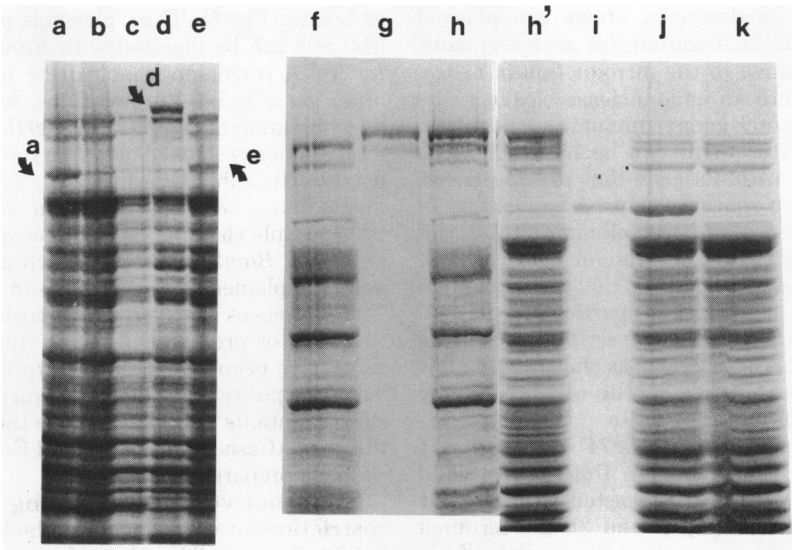

FIG. 3. *Gel analysis of hybrid β-galactosidase proteins. Whole-cell or partially purified proteins were run on 7.5% sodium dodecyl sulfate-polyacrylamide gels by the method of Laemmli (27). Cells were grown to the early stationary phase in LB medium and washed with 10 mM Tris (pH 7.4)-1 mM EDTA before being suspended in Laemmli sample buffer and boiled. Data from three separate gels (a through e, f through h, and h' through k) are shown. (a) Extracts of isopropyl-β-D-thiogalactopyranoside-induced E. coli Lac⁺ strain MC1020, showing the location of the 135,000-molecular weight wild-type β-galactosidase (arrow a). (b and c) Whole-cell extracts from uninduced MC1020, showing absence of the β-galactosidase band. (c) Diluted fivefold. (d) Whole-cell extract from E. coli strain MC1000 (Δlac) carrying a pMC874 plasmid derivative that contains the cloned BamHI fragment shown in track b of Fig. 2; a (hybrid) peptide larger than the β-galactosidase is seen (arrow d). (e) Whole-cell extract from MC1000 carrying the pMC874 plasmid that includes the BamHI fragment shown in track c of Fig. 2. A high-molecular-weight hybrid protein is seen (arrow e). (f) Extract from E. coli strain MC1000 carrying the pMC874 plasmid vector without an insert. (g) Protein obtained upon purification of hybrid β-galactosidase activity from the extract shown in h and h' (run in duplicate on different gels). The relative position of the peptide purified for β-galactosidase activity is the same as that of the new high-molecular-weight peptide seen in gels (lane d) of cells containing pMC874 derivatives that include the cloned BamHI fragment inserts. The size of the enzymatically active hybrid β-galactosidase is different from that of the wild-type purified β-galactosidase shown in lane i and of whole-cell extracts in lanes a and j. (k) Uninduced MC1020 cell extract (same as lane b).*

carrying only pMC874 and which migrated more slowly in the gel than the full-length β-galactosidase (lanes d and e). Purification of the β-galactosidase activity from extracts of one of these clones yielded a protein that migrated at the same position as the new peptide produced by the corresponding Lac⁺ clone (Fig. 3, lanes g and h). This result establishes that the high-molecular-weight peptide was made as a consequence of DNA insertion at the *Bam*HI site of plasmid pMC874. Since this vector lacked both transcription and translational start signals capable of yielding Lac expression, initiation of DNA and protein synthesis for the hybrid peptide necessarily occurred on the inserted DNA fragment.

Transcriptional fusion plasmid analogous to pMC874. Plasmid pMC874 can be used to identify and clone DNA fragments that initiate transcription and translation internally and yield both transcriptional and translational read-through into *lacZ*. For investigations of the relative effects of transcriptional and translational control signals in gene expression, it is useful to have an analogous plasmid that can select fragments containing only transcriptional promoters. For this purpose, plasmid pMC871 in which the *lac* promoter is replaced by a *trp* operon-derived segment of DNA that does not promote or block transcription, was constructed (Fig. 4; details in Materials and Methods). The *trp-lac* segment of this plasmid was originally derived from the W205 *trp-lac* fusion (33) and contains an intact *lacZ* gene, including a translation initiation site such that wild-type β-galactosidase is produced (38). Similar *lac* fusion plasmids for isolating endonuclease-generated DNA fragments that contain transcriptional promoters have been described in detail previously (11).

Fusion of *lac* gene segments of plasmids pMC874 and pMC871 into other vectors.

The constructs described above are plasmid cloning vectors that contain *lac* gene segments. As an alternative to the introduction of exogenous genes into an endonuclease cleavage site preceding the *lacZ* gene remnant in such vectors, the *lac* DNA segment can be inserted into a cleavage site within a gene that previously has been cloned on another vector. An advantage of this approach is that it can eliminate transcription that may extend into the *lac* gene segment from promoters located on the vector component of the plasmids; in particular, plasmid pMC871 has significant transcription from the vector into its *lac* segment as shown by its low level of *lac* expression in the absence of any DNA insert in its *Bam*HI site.

Whereas plasmids pMC874 and pMC871 yielded a *lac* fragment free from the plasmid replication region when digested with *Bam*HI plus *Sal*I or *Pst*I (Fig. 1 and 2), the terminal sequences at the two ends of the resulting fragment were different. Frequently, it is not convenient to join such fragments to termini generated by a single endonuclease cleavage within a gene. For this reason, we constructed the pMC931 and pMC903 derivatives of pMC874 and pMC871, respectively, that had *Bam*HI cleavage sites located after the *lac* genes as well as before (Fig. 5). Thus, plasmids pMC931 and pMC903 can be digested with *Bam*HI to yield *lac* DNA fragments that can be inserted into other genes in both orientations. Alternatively, a combination of *Bam*HI and *Bgl*II (which generates cohesive ends identical to those produced by *Bam*HI) can be used (Fig. 5). Such *lac* fragments have been inserted into genes within transposable element Tn*3* at cleavage sites generated by *Bam*HI or *Bcl*I (which also produce ends complementary to *Bam*HI) to fuse *lacZ* to Tn*3* sequences that serve as translation initiation sites or promoters or both (19). These fusions have been used to study expression of the Tn*3* transposition gene (19) and, furthermore, to isolate mutants that overproduce the product of this gene (Casadaban, Chou, and Cohen, manuscript in preparation).

***lac* fusion vectors containing additional restriction sites.** To facilitate the formation of hybrid β-galactosidase gene fusions with DNA fragments containing different types of termini, we constructed derivatives of pMC874 which had *Eco*RI and *Sma*I (*Xma*I) sites immediately adjacent to the *Bam*HI site for inserting DNA fragments. These plasmids were made by joining synthetic DNA sequences containing adjacent *Eco*RI, *Sma*I (*Xma*I), and *Bam*HI sites (3) on

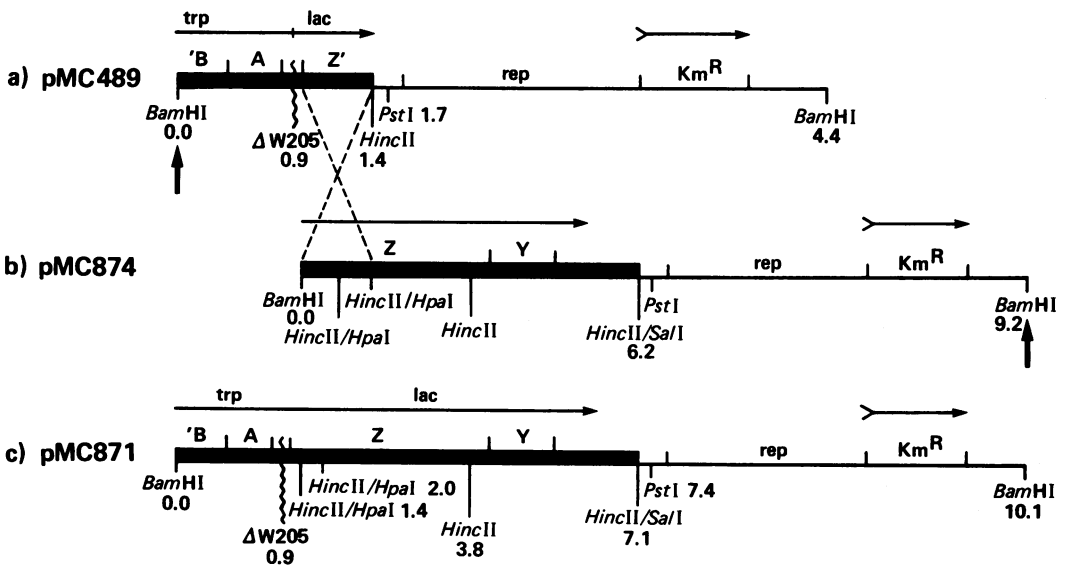

FIG. 4. *Construction of the pMC871 BamHI-lac transcriptional fusion vector. (a) pMC489 (11) contains a HincII-HpaI fragment of the W205 trp-lac fused operon (33) (heavy line) that ends at the HincII-HpaI site in the 146th amino acid codon of lacZ. The trp segment replaces the lac promoter and does not block transcription (33). This fragment was inserted between the BamHI and HincII cleavage sites of the pACYC177 vector with a synthetic BamHI oligonucleotide DNA linker (2) to reform the BamHI site, as described (11). (b and c) pMC871 was formed from pMC489 and pMC874 by homologous recombination (broken cross-lines) and by BamHI digestion (at arrows) and ligation as described in the text. These maps show the plasmids arbitrarily opened at their unique BamHI sites. Other designations are as in Fig. 1 and as described by Casadaban and Cohen (11).*

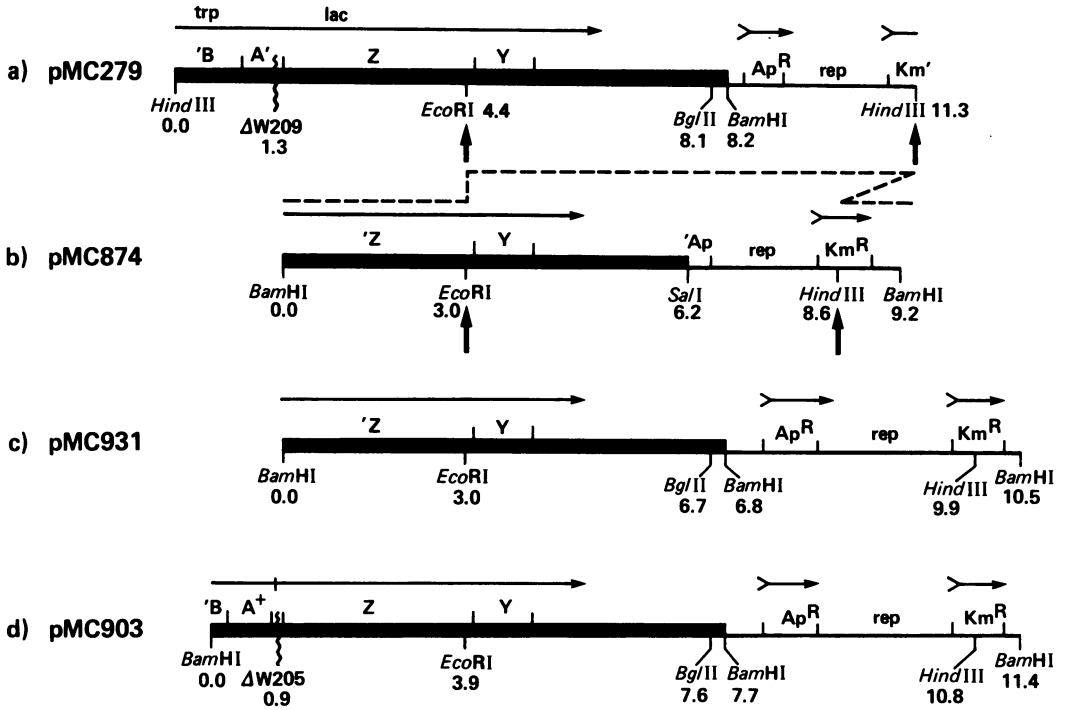

FIG. 5. *Structure of BamHI-lac fusion fragments lacking replicons. BamHI and BglII cleavage sites were introduced behind the lac genes of pMC874 and pMC871 by exchanging an EcoRI-HindIII fragment of each of these plasmids with the corresponding EcoRI-HindIII fragment of pMC279, as shown in (a) and (b). The resulting plasmids were pMC931 (c) (from pMC874) and pMC903 (d) (from pMC871). The arrows indicate the HindIII and EcoRI sites used for exchanging the endonuclease-generated fragments. pMC279 contains the lacZ and Y genes and has unique BglII and BamHI sites located after the lac genes. It was constructed by introducing a HindIII-BamHI fragment (thick line) containing the trp-lac fusion of plasmid pMC81 (11) into the pACYC177 plasmid vector that had been cleaved with both HindIII and BamHI endonucleases.*

the pBR322 plasmid to the *Bam*HI-*Sal*I *lac* fragment of pMC874, as described in Materials and Methods (Fig. 6). The pMC1396 plasmid retained the *Eco*RI site in *lacZ*, whereas in plasmid pMC1403 this additional *Eco*RI site was removed by mutation. Plasmid pMC1403 therefore had unique *Eco*RI, *Sma*I (*Xma*I), and *Bam*HI sites, which means that it incorporated into a single vector the ability to form hybrid β-galatosidase proteins by using cloned DNA fragments having termini generated by a wide variety of restriction endonucleases.

DISCUSSION

During the past several years, various types of gene fusions involving the *lac* genes of *E. coli* have been used for the production of peptides ordinarily foreign to this bacterial species (21, 23, 31, 35) and for the detection and analysis of DNA sequences that contain transcriptional regulatory signals (9, 10, 11, 14–17, 37). For foreign gene expression, the *lac* gene regulatory control region and the DNA segment encoding the amino-terminal end of its gene product have been fused at a conveniently located endonuclease cleavage site to exogenously derived DNA (31, 35) or to chemically synthesized DNA segments (21, 23). For genetic analysis, *lac* gene fusions have proved useful for detecting and isolating DNA fragments that contain transcriptional regulatory signals introduced upstream from the gene (11, 14–17, 37).

We have described here the construction of a new type of plasmid cloning vector for the fusion of the *lac* structural genes to transcriptional and translational regulatory signals of other genes. These plasmids have versatile *Bam*HI, *Eco*RI, and *Sma*I (*Xma*I) endonuclease cleavage sites. *Bam*HI cleaves the sequence G↓GATCC, yielding 5'-protruding GATC ends that are easily ligated to other *Bam*HI-generated termini or to complementary ends generated by a variety of restriction endonucleases including *Bgl*II (A↓GATCT), *Bcl*I (T↓GATCA), and *Sau*3A or *Mbo*I (↓GATC). *Eco*RI (G↓AATTC) generates

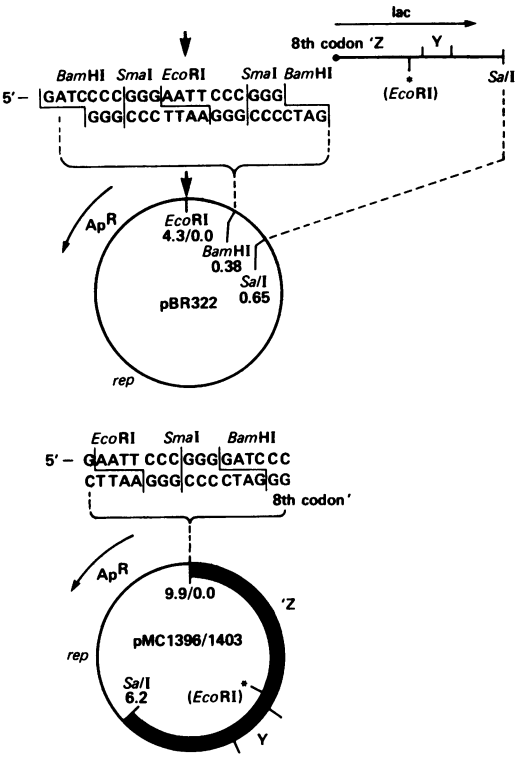

FIG. 6. *EcoRI-SmaI (XmaI)-BamHI site derivatives of the β-galactosidase fusion vector. The BamHI site by the eighth codon of lacZ was adapted to EcoRI and SmaI sites by using synthetic DNA linkers (3). (The SmaI site is also recognized by the XmaI endonuclease, which produces 5' protruding DNA ends.) A chemically synthesized BamHI-SmaI-EcoRI-SmaI-BamHI DNA sequence (ref. 2; sequence in figure) was first inserted in the BamHI site of pBR322 by L. Lau and R. Wu to form plasmid clone pLL10. This plasmid has two cleavage sites for EcoRI, BamHI and SmaI. One cleavage site for each endonuclease was removed by excising the DNA sequence between the two EcoRI sites (arrows) to form plasmid pMC1364, which has unique adjacent sites for the three endonucleases. Next, the BamHI-SalI lac segment from pMC874 was placed into pMC1364 to form pMC1396 (lower plasmid). pMC1396 thus contains a unique SmaI site adjacent to the BamHI site by lacZ, and this can be used to clone DNA fragments that yield hybrid β-galactosidases. pMC1396 does not have a unique EcoRI site since there is an EcoRI site in lacZ (labeled*). To make a plasmid similar to pMC1396 and lacking an EcoRI site in lacZ, we recombined a lacZ EcoRI site mutation (see text) onto pMC874 to form pMC1388. The BamHI-SalI lac fragment from this plasmid (equivalent to the pMC874 fragment, except for the EcoRI site) was then inserted into pMC1364 to form pMC1403, which is analogous to pMC1396 except for the EcoRI site in lacZ. pMC1403 can thus be used to clone DNA fragments at its EcoRI, SmaI, or BamHI site, yielding hybrid β-galactosidases in each instance.*

protruding AATT 5' ends which can be ligated to *Eco*RI- and *Eco*RI*- (N↓AATTN) (36) generated fragments. *Sma*I (CCC↓GGG) yields blunt ends which can be ligated to DNA fragments generated by endonucleases having many different sequence specificities and to DNA fragments made blunt-ended by other means. The fact that some of these endonucleases recognize a six-base-pair nucleotide sequence, whereas others recognize a four-base-pair sequence, provides flexibility by allowing cleavage of DNA either at multiple loci within a short segment or at relatively distant sites. Combinations of the six nucleotide-cleaving enzymes or partial digestions with four nucleotide-cleaving enzymes can be used to produce cleavages at intermediate distances if desired. This flexibility may also accomplish the production of functional *lac* fusions in instances in which translational reading of genes on *Bam*HI-generated fragments is out of frame with the β-galactosidase. *Bam*HI-, *Eco*RI-, and *Xma*I- (C↓CCGGG) cleaved vectors can also be joined to blunt-ended DNA either by using synthetic oligonucleotide linkers (2), by filling in the protruding ends with reverse transcriptase (24) or DNA polymerase (1, 39), or by removing the single-stranded projection, using a single-strand endonuclease such as S1 (42). Potentially, such manipulations allow linkages in different translational reading frames at the same endonuclease cleavage site. Moreover, even out-of-frame fusions that yield low levels of β-galactosidase can be detected by using appropriate indicator media.

The hybrid β-galactosidase gene fusions described here are potentially useful for a variety of purposes. The β-galactosidase activity of a hybrid protein can be used to assay expression of the gene fused to *lac* and to isolate mutants having altered expression (15). Genetic regulatory signals controlling promotion of transcription or initiation of translation of the original gene are applied in these fusions to the *lac* genes, as has been seen for all genes so far fused to *lacZ* in vivo (4, 9, 12, 22, 34, 40), and their effects can be measured by following *lac* expression. Functions attributable to a particular site or domain of the exogenously derived component of the hybrid protein (for example, a substrate-binding site) can be assigned to a position within the original protein by comparing hybrids of different sizes for the retention or loss of the site. This has been done for the operator-binding site of DNA binding of the *lacI* repressor protein (25, 34) and for membrane-binding sites for maltose transport proteins (5, 43, 44).

The approach described here also allows the construction of hybrid β-galactosidase genes that contain amino-terminal segments of genes derived from species other than *E. coli*. We have recently constructed such fusions in which expression of the *lacZ* gene in *E. coli* is mediated by transcriptional and translational signals contained on fragments of yeast chromosomal DNA (Casadaban et al., in preparation). Introduction of the yeast-*lac* fusion segments into yeast cells by transformation has resulted in expression of yeast-*E. coli* β-galactosidase hybrid proteins in *Saccharomyces cerevisiae*, using gene regulatory signals of the yeast. *lac* gene translational fusions may similarly permit the detection and cloning of transcriptional and translational signals from "higher" eucaryotic organisms, using production of enzymatically active β-galactosidase as a direct assay.

ACKNOWLEDGMENTS

We thank L. Lau and R. Wu for the gift of plasmid pLL10, K. Bertrand and W. Reznikoff for the gift of λRS205, and our many colleagues who have contributed advice.

These studies were supported by Public Health Service grant AI 08619 from the National Institutes of Health and in part by grant VC-139D from the American Cancer Society and Public Health Service grant GM 26355 from the National Institutes of Health to S.N.C. M.J.C. was a postdoctoral research fellow of the National Institutes of Health.

LITERATURE CITED

1. **Backman, K., M. Ptashne, and W. Gilbert.** 1976. Construction of plasmids carrying the CI gene of bacteriophage λ. Proc. Natl. Acad. Sci. U.S.A. **73**:4174-4178.
2. **Bahl, C., K. Marians, R. Wu, J. Strawinsky, and S. Narang.** 1976. A general method for inserting specific DNA sequences into cloning vehicles. Gene **1**:81-92.
3. **Bahl, C., R. Wu, R. Brousseau, A. Sood, H. Hsiung, and S. Narang.** 1978. Chemical synthesis of versatile adaptors for molecular cloning. Biochem. Biophys. Res. Commun. **81**:695-703.
4. **Bassford, P., J. Beckwith, M. Berman, E. Brinkman, M. Casadaban, L. Guarente, I. Saint-Girons, A. Sarthy, M. Schwartz, H. Shuman, and T. Silhavy.** 1978. Genetic fusions of the *lac* operon: a new approach to the study of biological processes, p. 245-261. *In* J. Miller and W. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
5. **Bassford, P., T. Silhavy, and J. Beckwith.** 1979. Use of gene fusion to study secretion of maltose-binding protein into *Escherichia coli* periplasm. J. Bacteriol. **139**:19-31.
6. **Beckwith, J.** 1978. lac: the genetic system, p. 11-30. *In* J. Miller and W. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
7. **Bolivar, F., R. Rodriguez, M. Betlach, and H. Boyer.** 1977. Construction and characterization of new cloning vehicles. I. Ampicillin-resistant derivatives of the plasmid pMB9. Gene **2**:75-93.
8. **Casadaban, M.** 1975. Fusion of the *E. coli lac* genes to the *ara* promoter: a general technique using bacteriophage Mu-1 insertions. Proc. Natl. Acad. Sci. U.S.A. **72**:809-813.
9. **Casadaban, M.** 1976. Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. J. Mol. Biol. **104**:541-555.
10. **Casadaban, M., 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. U.S.A. **76**: 4530-4533.
11. **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.
12. **Casadaban, M., T. Silhavy, M. Berman, H. Shuman, A. Sarthy, and J. Beckwith.** 1977. Construction and use of gene fusions directed by bacteriophage Mu insertions, p. 531-535. *In* H. Bukhari, J. Shapiro, and S. Adhya (ed.), DNA insertion elements, plasmids, and episomes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Chang, A., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the p15A cryptic miniplasmid. J. Bacteriol. **134**:1141-1156.
14. **Chou, J., M. Casadaban, P. Lemaux, and S. N. Cohen.** 1979. Identification and characterization of a self-regulated repressor of translocation of the Tn3 element. Proc. Natl. Acad. Sci. U.S.A. **76**:4020-4024.
15. **Chou, J., P. Lemaux, M. Casadaban, and S. N. Cohen.** 1979. Transposition protein of Tn3: identification and characterization of an essential repressor-controlled gene product. Nature (London) **282**:801-806.
16. **Cohen, S. N., F. Cabello, M. Casadaban, A. C. Y. Chang, and K. Timmis.** 1977. DNA cloning and plasmid biology, p. 35-38. *In* W. Scott and R. Werner (ed.), Molecular cloning of recombinant DNA, Miami Winter Symposia, vol. 13. Academic Press, Inc., New York.
17. **Cohen, S. N., M. Casadaban, J. Chou, and C. Tu.** 1979. Studies of the specificity and control of transposition of the Tn3 element. Cold Spring Harbor Symp. Quant. Biol. **43**:1247-1255.
18. **Cohen, S. N., A. Chang, H. Boyer, and R. Helling.** 1973. Construction of biologically functional bacterial plasmids *in vitro*. Proc. Natl. Acad. Sci. U.S.A. **70**: 3240-3244.
19. **Cohen, S. N., A. Chang, and L. Hsu.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *E. coli* by R-factor DNA. Proc. Natl. Acad. Sci. U.S.A. **69**:2110-2114.
20. **Fowler, A., and I. Zabin.** 1977. The amino acid sequence of β-galactosidase of *E. coli*. Proc. Natl. Acad. Sci. U.S.A. **74**:1507-1510.
21. **Goeddel, D., D. Kleid, F. Bolivar, H. Heyneker, D. Yansura, R. Crea, T. Hirose, A. Kraszewski, K. Itakura, and A. Riggs.** 1979. Expression in *Escherichia coli* of chemically synthesized genes for human insulin. Proc. Natl. Acad. Sci. U.S.A. **76**:106-110.
22. **Heidecker, G., and B. Muller-Hill.** 1977. Synthetic multifunctional proteins. Mol. Gen. Genet. **155**:301-307.
23. **Itakura, K., T. Hirose, R. Crea, A. Riggs, H. Heyneker, F. Bolivar, and H. Boyer.** 1977. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. Science **198**:1056-1063.
24. **Kacian, D., and S. Spiegelman.** 1972. Purification and detection of reverse transcriptase in viruses and cells. Methods Enzymol. **29**:150-173.
25. **Kania, J., and B. Muller-Hill.** 1977. Construction, isolation and implications of repressor-galactosidase: β-galactosidase hybrid molecules. Eur. J. Biochem. **79**: 381-386.
26. **Kupersztoch, Y., and D. Helinski.** 1973. A catenated DNA molecule as an intermediate in the replication of the resistance transfer factor R6K in *Escherichia coli*. Biochem. Biophys. Res. Commun. **54**:1451-1459.
27. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680-685.
28. **Langley, K., M. Villarejo, A. Fowler, P. Zamenhof,**

and I. Zabin. 1975. Molecular basis of β-galactosidase α-complementation. Proc. Natl. Acad. Sci. U.S.A. **72:** 1254-1257.

29. **Maniatis, T., A. Jeffrey, and H. Van de Sande.** 1975. Chain length determination of small double- and single-stranded DNA molecules by polyacrylamide gel electrophoresis. Biochemistry **14:**3787-3794.
30. **Maxam, A., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74:**560-564.
31. **Mercereau-Puijalon, O., A. Royal, B. Cami, A. Garapin, A. Krust, F. Gannon, and P. Kourilsky.** 1978. Synthesis of an ovalbumin-like protein by *Escherichia coli* K12 harboring a recombinant plasmid. Nature (London) **275:**505-510.
32. **Miller, J.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. **Mitchell, D., W. Reznikoff, and J. Beckwith.** 1975. Genetic fusions defining *trp* and *lac* operon regulatory elements. J. Mol. Biol. **93:**331-350.
34. **Muller-Hill, B., and J. Kania.** 1974. Lac repressor can be fused to β-galactosidase. Nature (London) **249:**561-562.
35. **O'Farrell, P., B. Polisky, and D. Gelfand.** 1978. Regulated expression by readthrough translation from a plasmid-encoded β-galactosidase. J. Bacteriol. **134:**645-654.
36. **Polisky, B., P. Greene, D. Garfin, B. McCarthy, H. Goodman, and H. Boyer.** 1975. The specificity of substrate recognition by the *Eco*RI restriction endonuclease. Proc. Natl. Acad. Sci. U.S.A. **72:**3310-3315.
37. **Ptashne, M., A. Jeffrey, A. Johnson, R. Maurer, B. Meyer, C. Pabo, T. Roberts, and R. Sauer.** 1980. How the repressor and *cro* work. Cell **19:**1-11.
38. **Reznikoff, W., C. Michels, T. Cooper, A. Silverstone, and B. Magasanik.** 1974. Inhibition of *lacZ* gene translation initiation in *trp-lac* fusion strains. J. Bacteriol. **117:**1231-1239.
39. **Richardson, C., C. Schildkraut, H. Aposhian, and A. Kornberg.** 1964. Enzymatic synthesis of deoxyribonucleic acid. J. Biol. Chem. **239:**222-232.
40. **Sarthy, A., A. Fowler, I. Zabin, and J. Beckwith.** 1979. Use of gene fusions to determine a partial signal sequence of alkaline phosphatase. J. Bacteriol. **139:** 932-939.
41. **Sharp, P., B. Sugden, and J. Sambrook.** 1975. Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. Biochemistry **12:** 3055-3063.
42. **Shenk, T., C. Rhodes, P. Rigby, and P. Berg.** 1975. Biochemical method for mapping mutational alterations in DNA with S1 nuclease: the location of deletions and temperature-sensitive mutations in SV40. Proc. Natl. Acad. Sci. U.S.A. **72:**989-993.
43. **Silhavy, T., M. Casadaban, H. Shuman, and J. Beckwith.** 1976. Conversion of β-galactosidase to a membrane-bound state by gene fusion. Proc. Natl. Acad. Sci. U.S.A. **73:**3423-3427.
44. **Silhavy, T., H. Shuman, J. Beckwith, and M. Schwartz.** 1977. Use of gene fusions to study outer membrane protein localization in *E. coli*. Proc. Natl. Acad. Sci. U.S.A. **74:**5411-5415.
45. **Sutcliffe, G.** 1979. Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. **43:**77-90.
46. **Timmis, K., F. Cabello, and S. N. Cohen.** 1978. Cloning and characterization of *Eco*RI and *Hin*dIII resistance plasmids R6-5 and R6. Mol. Gen. Genet. **162:**121-137.
47. **Welply, J., A. Fowler, J. Beckwith, and I. Zabin.** 1980. Position of early nonsense and deletion mutations in *lacZ*. J. Bacteriol. **142:**732-734.