JOURNAL OF BACTERIOLOGY, May 1978, p. 645-654
0021-9193/78/0134-0645$2.00/0

Vol. 134, No. 2
Printed in U.S.A.

Regulated Expression by Readthrough Translation from a Plasmid-Encoded β-Galactosidase

PATRICK H. O'FARRELL,[1] BARRY POLISKY,[2] AND DAVID H. GELFAND[3]*

Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143;[1] Program in Molecular, Cellular and Developmental Biology, Indiana University, Bloomington, Indiana 47401;[2] and Cetus Corporation, 600 Bancroft Way, Berkeley, California 94710[3]

Received for publication 28 October 1977

We have characterized expression of β-galactosidase from a plasmid cloning vehicle, pBGP120, which carries most of the *lacZ* gene and contains a single *Eco*RI site near the end of *lacZ*. In addition, we have examined expression of heterologous DNA inserted at the position of the *Eco*RI site. The *Eco*RI site was shown to be within the sequence coding for β-galactosidase, and its precise location and phase were deduced. Insertion of heterologous *Eco*RI-generated DNA fragments altered the molecular weight of the plasmid-encoded β-galactosidase polypeptide. Those insertions that were in the correct phase were expressed at a high level as a fused protein. The different forms of β-galactosidase polypeptides produced by various hybrid plasmids were all stable proteins. The level of expression of the plasmid-encoded β-galactosidase was several times higher than maximal expression of chromosome-encoded β-galactosidase, suggesting that expression is proportional to gene copy number. The expression of the plasmid *lacZ* gene was controlled by cyclic AMP. When grown in a *cya* strain (DG74), expression was dependent on exogenous cyclic AMP. Although in normal strains there was insufficient *lac* repressor to inactivate all copies of the plasmid, repressor regulation was restored when the plasmid was grown in a strain (M96) that overproduces the *lac* repressor.

A multicopy plasmid cloning vehicle permitting controlled expression of inserted heterologous DNA fragments has been constructed (17). This plasmid, pBGP120, contains the *lac* promoter and operator, the structural gene for β-galactosidase (*lacZ*), and a single *Eco*RI restriction endonuclease site "downstream" from the *lac* promoter. The introduction of well-characterized bacterial regulatory signals into a cloning vehicle was intended for studies of the expression of eucaryotic genetic information in bacteria. Indeed, when *Xenopus laevis* 28S ribosomal DNA (rDNA) sequences were inserted at the *Eco*RI site of pBGP120, transcription of these sequences was modulated by elements affecting *lac* operon expression (17).

To further assess the usefulness of pBGP120 as a cloning vehicle, we have studied *lac* operator-controlled synthesis of proteins in the vehicle and hybrid plasmids. The plasmid-specified β-galactosidase polypeptide is shown to be defective structurally and enzymatically, and its molecular weight depends on the nature of the DNA sequences inserted at the *Eco*RI site of pBGP120. We also show that inserted *Eco*RI fragments of heterologous DNA can be expressed as stable products of readthrough translation. Finally, the location of the *Eco*RI recognition site in *lacZ* is discussed.

MATERIALS AND METHODS

Bacterial strains and plasmids. The strains used in this study are all derivatives of *Escherichia coli* K-12. Their designations and characteristics are listed in Table 1. Strain DG73 contained an F factor derived from strain W3140 (17). The strain was cured of F by growth with acridine orange (11), giving rise to strain DG75. The construction of plasmids pBGP120, pBGP123, and pBGP124 has been described (17).

Media and culture methods. Minimal medium was medium E of Vogel and Bonner (21) or M9 (11) supplemented with 0.5% glucose, 10 μg of thiamine per ml, and 50 to 100 μg of required amino acids per ml. When indicated, minimal medium was supplemented with Norit-treated vitamin-free Casamino Acids (Difco Laboratories) to 0.25%. β-Lactose (Eastman Organic Chemicals) was added to 0.2% when necessary. L broth contained 1% tryptone (Difco), 0.5% yeast extract (Difco), 0.2% glucose, and 0.5% NaCl. Drug-selective media contained 20 to 30 μg of ampicillin (Bristol Laboratories) per ml. Media were solidified with 1.5% agar (Difco). Lactose indicator plates were either MacConkey agar base (Difco) supplemented with lactose (1%) or 5-bromo-4-chloro-3-indolyl-β-D-galactosidase (X-gal) indicator plates (11). X-gal is a colorless, noninducing substrate for β-galactosidase.

TABLE 1. *Bacterial strains and plasmids*

Strain/plasmid	Genotype/relevant characteristic	Source
Bacterium		
DG72	*hsdS1 leu-6 proA2 ara-14 galK2 xyl-5 mtl-1 rpsL20 thi-1 supE44* λ^- F^-	Laboratory stock
DG73	As in DG72 *lacZ39* Pro^+ F^+	Laboratory stock
DG74	As in DG73 *cya-283*	P1 transduction from LS680 and selection for Cya^-
DG75	As in DG73 F^-	Acridine orange cured (11)
M96	*ara* Δ*lac thi* [λcI857*St*68h80d*lacI*(Sq) *lacZ*u118]	A. Riggs
ΔLW	*his* Δ*lac* F^-	D. Zipser
Plasmid		
pBGP120	ColE1-derived plasmid, Ap^r *lacO*$^+$ *lacP*$^+$, major portion of *lacZ* gene, single "downstream" *Eco*RI site	(17, this paper)
pBGP123	pBGP120 containing *X. laevis* 28S rDNA	(17)
pBGP124	As in pBGP123 with *X. laevis* DNA in opposite orientation	(17)

When hydrolyzed, a deep blue color is formed, permitting detection of β-galactosidase activity in cells grown in the presence of glucose or glycerol as the sole carbon source. Liquid cultures were grown at 37°C in a New Brunswick gyratory shaker, and densities were monitored by observing the increase in absorbance at 660 nm. Plasmid DNA was isolated after chloramphenicol amplification of plasmid copy number, detergent lysis of lysozyme protoplasts, and phenol extraction (1). Covalently closed supercoiled DNA was isolated after propidium diiodide (Calbiochem) CsCl equilibrium ultracentrifugation (6). Propidium diiodide was removed by repeated extraction with CsCl-saturated isopropanol.

Cloning methods. DNA ligation (5), agarose slab gel electrophoresis (4), transformation of bacteria (5), and a rapid screening procedure for covalently closed circular DNA from individual cultures (10) were essentially as described. Identity and orientation of cloned DNA fragments were determined by using methods analogous to those previously published (17).

Labeling, sample preparation, and electrophoresis of proteins. *E. coli* strains were grown at 30°C in minimal M9 glucose media with appropriate amino acid supplements added to 100 μg/ml and were labeled for 5 or 10 min in 5 to 25 μCi of [^{35}S]methionine per ml (about 400 Ci/mmol). When required, cyclic AMP was added to 3×10^{-2} M, isopropylthiogalactoside (IPTG) was added to 2.5×10^{-4} M, and the cells were incubated for 15 min before labeling. After labeling, cells were chilled on ice, collected by centrifugation, and prepared for electrophoresis as described previously (14).

Two-dimensional gels were prepared, and proteins were subjected to electrophoresis as described previously (14, 15). In all cases, the isoelectric focusing gel contained 1.6% Ampholines (pH 5 to 7) and 0.4% Ampholines (pH 3.5 to 10) (LKB) and was focused for 6,400 V-h. The acrylamide concentrations of the second-dimension sodium dodecyl sulfate (SDS) gels are given in the figure legends. SDS-gel electrophoresis was performed by the procedure of Laemmli (8, 9) adapted for slab gels (14).

β-Galactosidase assays. β-Galactosidase was assayed as described (11) with or without induction by IPTG (2.5×10^{-4} M), with or without added cyclic AMP (30 mM).

Chemicals. Chemicals were obtained commercially and used without further purification. Restriction endonucleases *Bam*HI and *Hin*dIII and T4 DNA ligase were purchased from New England Biolabs. Restriction endonuclease *Eco*RI was a generous gift from Lou DeGennaro. Reagents for casting acrylamide gels were purchased from Bio-Rad Laboratories. Other reagents for acrylamide gel electrophoresis were from sources previously described (14). Agarose (ME) was obtained from Marine Colloids.

Safety procedures. These experiments were conducted in a P2 facility by using precautions outlined in the National Institutes of Health "Guidelines for Research Involving Recombinant DNA Molecules" (13). Liquid cultures were treated with Wescodyne (West Chemical Products, Inc.) and DNA solutions with bleach before disposal. Agar plates containing transformed bacteria were autoclaved in polypropylene "Disposabags" (Dolby Scientific).

RESULTS

Plasmid-encoded β-galactosidase. The plasmid pBGP120 contains a large portion of a 4.4×10^6-dalton *Eco*RI fragment containing *lacZ* sequences derived from λp*lac*5 (17). λp*lac*5 carries an intact β-galactosidase gene (18). We compared the electrophoretic properties of the β-galactosidase polypeptide encoded by λp*lac*5 and that produced by the plasmid-bearing strain, DG73(pBGP120). Figure 1A shows a portion of a two-dimensional electrophoretic separation of proteins from strain DG73 lysogenized with λp*lac*5. The lysogen was heat induced, and β-galactosidase was subsequently induced by addition of IPTG in the presence of cyclic AMP. The arrow (Fig. 1A) indicates the spot shown to correspond to the β-galactosidase monomer by three criteria: (i) comigration with purified monomer (7), (ii) inducibility by IPTG (7, and Fig. 3), and (iii) absence in strains with deletions of

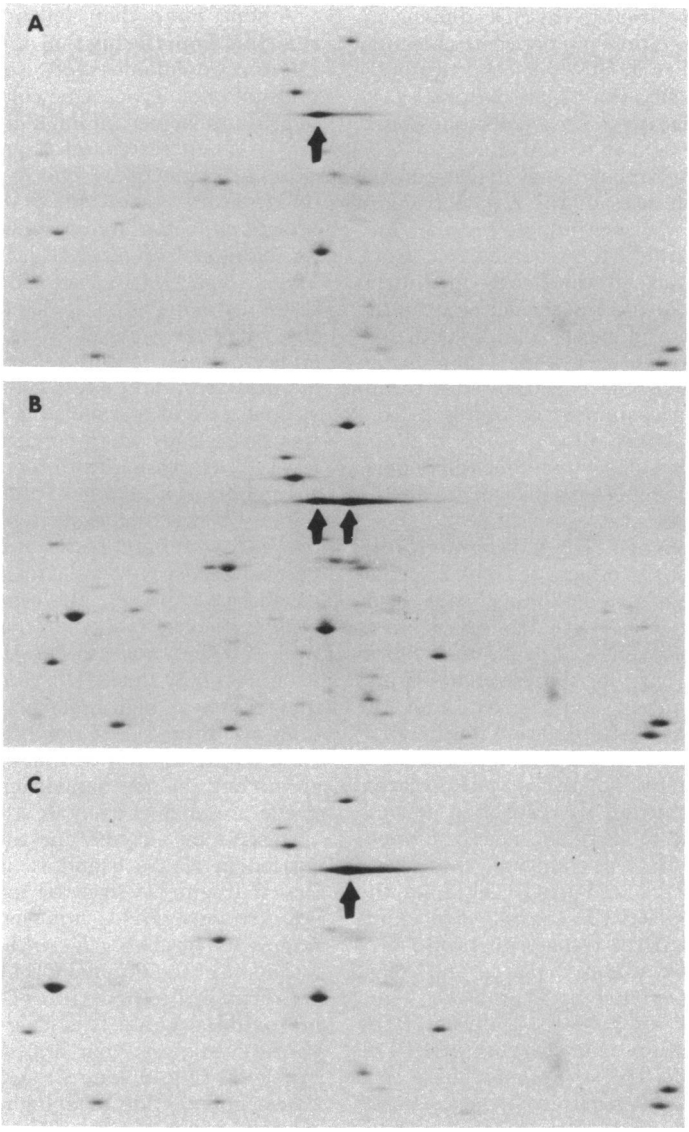

FIG. 1. *Comparison of the electrophoretic properties of wild-type and plasmid-encoded β-galactosidase monomers. Strains DG73(pBGP120) and a heat-induced λplac5 lysogen of DG73 were treated with cyclic AMP and IPTG for 15 min and labeled with [^{35}S]methionine for 5 min as described in the text. Extracts were prepared and subjected to electrophoresis in two dimensions. The horizontal dimension is isoelectric focusing with the basic side oriented toward the left. The vertical dimension is SDS-gel electrophoresis on a 10% acrylamide gel, and the high-molecular-weight polypeptides are toward the top. The panels shown are from portions of 20-h autoradiograms of the separations derived from λplac5 lysogen of DG73 (A), DG73(pBGP120) (C), and a mixture of the two samples (B). The arrow in (A) and the left arrow in (B) indicate the position of wild-type β-galactosidase monomers, whereas the arrow in (C) and the right arrow in (B) indicate the position of the plasmid-encoded β-galactosidase-related polypeptides.*

lacZ. The arrow in Fig. 1C indicates the position of the comparable polypeptide made by strain DG73 carrying pBGP120. Figure 1B shows a mixture of the samples from the lysogen and the transformant. The mixture clearly shows that the *lacZ* polypeptide encoded by the plasmid gene differs from the normal β-galactosidase monomer in isoelectric point, whereas no ob-

vious change is apparent in the SDS dimension. From the known relationship between molecular weight and mobility in SDS gels, the two polypeptides differ by less than 1,000 daltons.

These data suggest the presence of an *Eco*RI endonuclease site near the distal end of the *lacZ* gene. Because any translational and transcriptional termination sites of the β-galactosidase gene originally in λp*lac*5 would be consequently absent in the plasmid, we presumed that RNA polymerase read through the *Eco*RI endonuclease site in pBGP120 and transcribed some of the adjacent plasmid sequences. The defective β-galactosidase polypeptides would then terminate when the ribosome encountered a termination site in the sequences transcribed from adjacent plasmid DNA. This predicts that the size of the β-galactosidase monomer will depend on the sequence of DNA that is distal to the *Eco*RI endonuclease site.

Insertion of *Eco*RI DNA fragments in pBGP120. To test the effects of DNA insertion at the *Eco*RI site of pBGP120 on β-galactosidase, we constructed (17) pBGP120 derivatives containing *Xenopus* rDNA. The pSC101 hybrid plasmid CD4 (12), carrying the *Xenopus* 18S and 28S rDNA sequences, was cleaved with the *Eco*RI endonuclease. After ligation to pBGP120 and transformation of strain DG73, a number of clones carrying hybrid plasmids were isolated. Two were characterized by restriction enzyme analysis and found to carry 28S rDNA in opposite orientations (data not shown); these were designated pBGP123 and pBGP124. Like the parental plasmid, pBGP120, the hybrid plasmids pBGP123 and pBGP124 transform strain DG73 to Lac^+ on lactose minimal plates and X-gal indicator plates incubated at 30 or 42°C.

Figure 2 shows an SDS-gel analysis of the polypeptides made in cells carrying pBGP120, derivatives of pBGP120, or λp*lac*5. These cells were induced with IPTG in the presence of cyclic AMP and labeled with [^{35}S]methionine, and total polypeptides were separated by SDS-gel electrophoresis. The position of the wild-type β-galactosidase is indicated by the arrow. Lanes A and B show the polypeptides from the λp*lac*5 lysogen and heat-induced lysogen, respectively. β-Galactosidase monomer produced by pBGP120 is seen as a band comigrating with the wild-type enzyme (lane C). The β-galactosidase polypeptides specified by pBGP123 and pBGP124 are larger-than-normal β-galactosidase monomers by about 6,000 and 2,000 daltons, respectively (lanes E and F). No band comigrates with the wild-type β-galactosidase. This result shows that the size of the β-galactosidase monomer synthesized depends on the orientation of the inserted DNA.

Among more than 400 other transformants resulting from the ligation of pBGP120 and CD4 DNAs, two colonies were isolated that showed unusual phenotypes. One colony (clone 8) failed to grow on lactose minimal plates at either 30 or 42°C. When tested on X-gal indicator plates containing glucose, clone 8 was Lac^+ at both temperatures, showing that the absence of growth on lactose minimal plates was not due to an inability to hydrolyze lactose. As shown in Fig. 2 (lane G), β-galactosidase monomer produced by clone 8 was 28,000 daltons larger than the wild-type subunit. It is possible that the additional polypeptide sequence fused to β-galactosidase in clone 8 resulted in a hybrid protein still capable of metabolizing lactose, but which was deleterious when conditions required high levels of expression from the *lac* promoter.

In contrast to clone 8, another transformant (clone 12) was temperature sensitive for growth on lactose minimal plates and displayed a temperature-sensitive β-galactosidase activity on X-gal indicator plates. The temperature-sensitive β-galactosidase monomer specified by clone 12 was 48,000 daltons larger than the wild-type subunit (Fig. 2, lane D). It should be emphasized that the new high-molecular-weight polypeptides are altered *lacZ* polypeptides because the band comigrating with the wild-type subunit is absent and the new band is inducible with IPTG in the presence of cyclic AMP.

Restriction enzyme analysis of the plasmids carried in clones 8 and 12 indicated that the *Eco*RI fragments inserted into pBGP120 were not *Xenopus* rDNA, but were unknown DNA fragments (probably *E. coli*) present as a contaminant of the original CD4 DNA preparation.

Control of expression of the cloned β-galactosidase gene. It is clear from the electrophoretic analysis that transformants carrying pBGP120 (Fig. 2, lane C) synthesize β-galactosidase polypeptide at a higher rate than cells carrying a single copy of the *lacZ* gene (Fig. 2, lane A). However, the induced level of β-galactosidase enzymatic activity is low in the transformants (Table 2). Analysis of Coomassie brilliant blue-stained gels showed a high level of β-galactosidase polypeptide in strain DG73(pBGP120) (data not shown), indicating that the low level of enzymatic activity was not due to rapid degradation of the plasmid-encoded polypeptide. Thus, the β-galactosidase-like polypeptide encoded by pBGP120 has abnormally low enzymatic activity in strain DG73. In fact, a strain with a complete *lacZ* deletion (ΔLW), which was transformed with pBGP120, gave no β-galactosidase-positive transformants on X-gal ampicillin plates. Plasmid DNA isolated from these transformants was capable of transforming

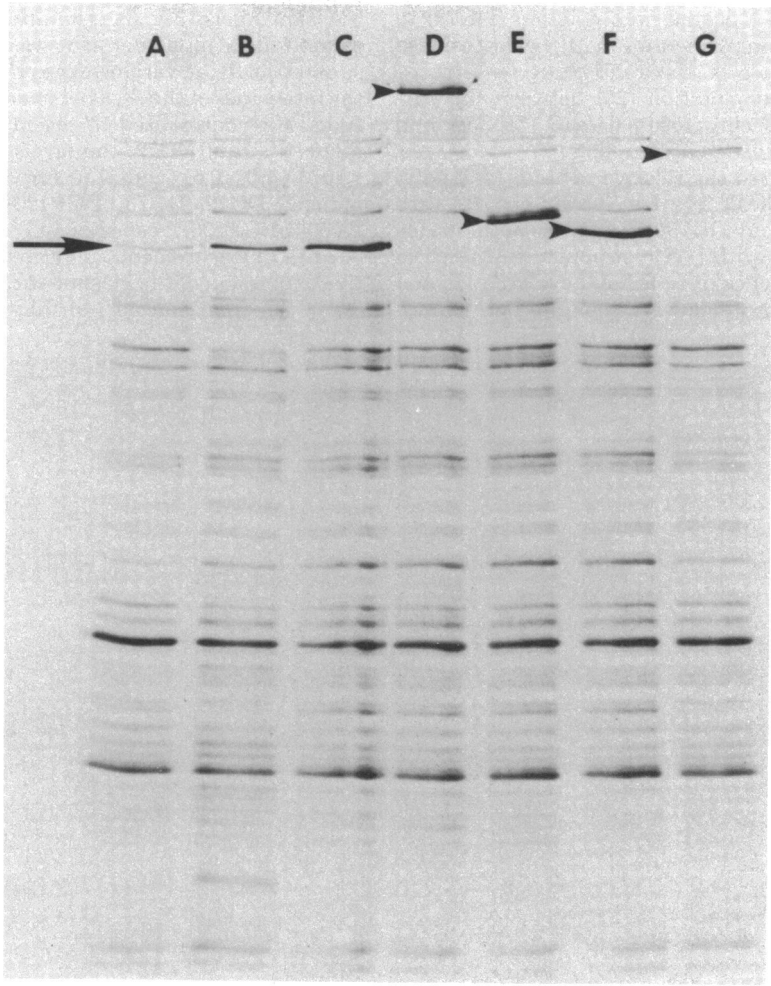

FIG. 2. *SDS-gel electrophoretic comparison of the sizes of β-galactosidase monomers encoded by λplac5, pBGP120, and hybrid derivatives of pBGP120. Cells were induced with IPTG in the presence of cyclic AMP and labeled with [^{35}S]methionine as described in the legend to Fig. 1. The samples (A) λplac5 lysogen of DG73, (B) a heat-induced λplac5 lysogen, (C) DG73(pBGP120), (D) clone 12 (a Lac(Ts) transformant of DG73), (E) DG73(pBGP123), (F) DG73(pBGP124), and (G) clone 8 (a Lac$^-$ transformant of DG73) were subjected to electrophoresis on a 7.5% acrylamide gel. An autoradiogram of the ^{35}S-labeled polypeptides is shown. The arrow indicates the position of the wild-type β-galactosidase monomer, and the arrowheads indicate the positions of the β-galactosidase-related polypeptides encoded by the hybrid plasmids.*

TABLE 2. *Abnormal β-galactosidase activity and inducibility in DG73(pBGP120)*

Strain	Relative β-galactosidase activity[a] with			
	Nothing	Cyclic AMP	IPTG	Cyclic AMP and IPTG
DG72 Lac$^+$	2.7×10^{-4}	3.6×10^{-4}	1	9.6
DG73(pBGP120) Lac$^+$	3.3×10^{-3}	3.7×10^{-2}	4.6×10^{-2}	0.5
DG73 Lac$^-$	5.9×10^{-7}	6.3×10^{-7}	5.8×10^{-7}	6.1×10^{-7}

[a] Enzymatic activities were measured 45 min after addition of cyclic AMP and/or IPTG to cells growing in M9 glucose medium at 37°C. Enzymatic activity was measured by the method of Miller (11), and data were normalized to the activity in IPTG-induced DG72 (1,300 nmol of *o*-nitrophenol per min per mg of protein). Cyclic AMP and IPTG were added to a final concentration of 30 mM and 2.5×10^{-4} M, respectively.

a strain with a partial *lacZ* deletion (DG73) to Lac$^+$. This complementation of β-galactosidase in strain DG73 is apparently due to intracistronic complementation (20) between the partially deleted chromosomal *lacZ* (*lacZ39*) and the plasmid *lacZ* gene products.

We compared the relative rates of *lacZ* polypeptide synthesis in a Lac$^+$ strain (DG72) with that of a strain DG73(pBGP120). Both strains were labeled with [^{35}S]methionine in the presence and absence of IPTG and cyclic AMP, and the labeled proteins were separated on two-dimensional gels (Fig. 3). The intensity of the β-galactosidase monomer spot was analyzed densitometrically at various exposure times. When the intensities of the β-galactosidase polypeptide spots were normalized to the intensity seen in induced strain DG72, the levels detected were $<10^{-3}$, 1.0, 0.5, and 7.0 for strains DG73, induced DG72, DG73(pBGP120), and induced DG73(pBGP120), respectively.

These measurements indicate that constitutive expression of the plasmid *lacZ* gene is about 8% of maximal and 50% of fully induced wild-

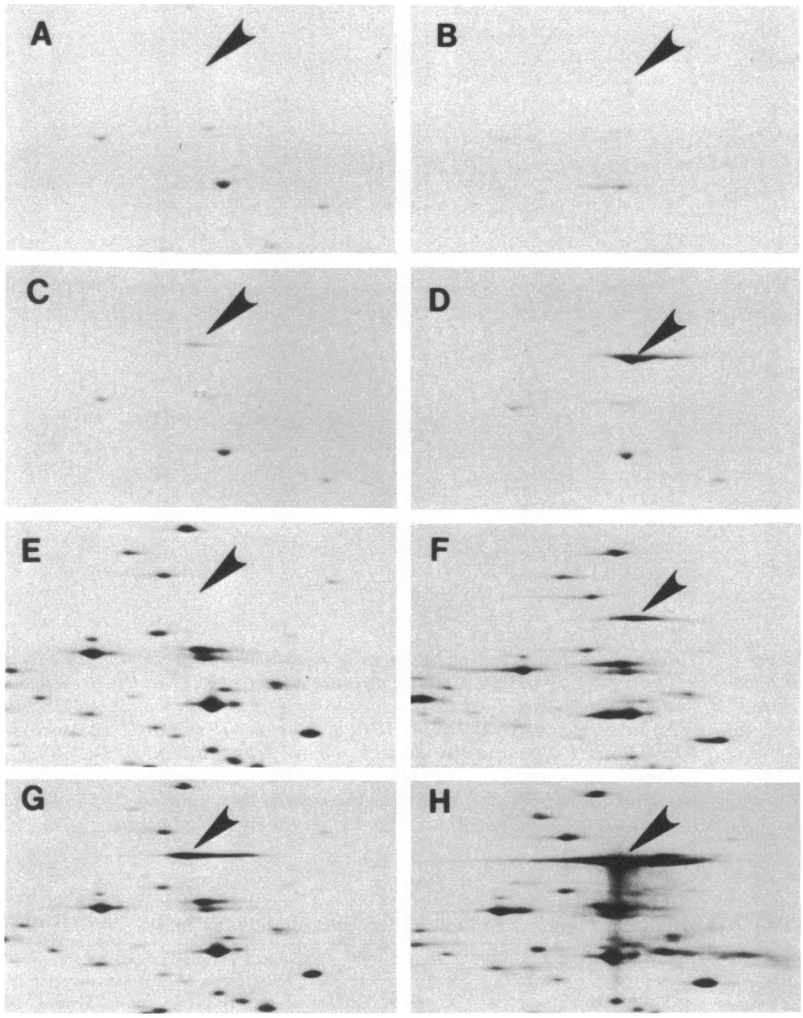

FIG. 3. *Electrophoretic measurement of the basal and induced rates of β-galactosidase synthesis in wild-type and pBGP120-transformed cells. Strains DG72 and DG73(pBGP120) were grown in minimal glucose media and labeled for 10 min with [^{35}S]methionine directly or after induction with cyclic AMP and IPTG. The samples (A and E) DG72, (B and F) DG73(pBGP120), (C and G) induced DG72, and (D and H) induced DG73(pBGP120) were subjected to electrophoresis in two dimensions. The same amount of total incorporated label was subjected to electrophoresis in each case, and 1.1-h exposures (A through D) and 22-h exposures (E through H) are shown.*

type expression. Enzymatic assays of β-galactosidase (Table 2) also indicated substantial constitutive expression in strain DG73(pBGP120). One obvious explanation for this observation is that the number of *lac* operator sites approaches or exceeds the number of *lac* repressor molecules so that some β-galactosidase genes are active even in the absence of inducer. The normal number of *lac* repressor molecules per cell is about 10 (3), whereas the plasmid copy number for Coli E1 derivatives normally exceeds 10 (19). To test whether the plasmid *lacZ* gene could be placed under repressor control, we transformed strain M96 with pBGP120. Strain M96 overproduces the *lac* repressor 50-fold due to a mutation in the *lacI* gene. Transformants plated on X-gal indicator plates were negative for β-galactosidase. However, when IPTG was included in the plates, they were positive, indicating that β-galactosidase synthesis is under repressor control.

We measured the rates of synthesis of β-galactosidase-related polypeptides in strain M96 transformed with pBGP120, pBGP123, or pBGP124. Figure 4 shows this electrophoretic analysis. The bands corresponding to the plasmid-encoded β-galactosidase monomer are inducible, demonstrating that expression remains dependent on the activity of the *lac* promoter. As seen in lane G of Fig. 4, some of these transformants show residual constitutive *lacZ* expression, although the *lac* repressor level should be sufficient to reduce expression below detectable levels. This strain produced *lac* constitutive segregants at a high frequency when plated on X-gal plates. This could account for the uninduced expression seen in Fig. 4.

The activity of the *lac* promoter shows a stringent requirement for cyclic AMP in adenyl cyclase-deficient cells (16). A Cya^- derivative of strain DG73 (DG74) was transformed by pBGP120 or several hybrid plasmids. These transformants showed no β-galactosidase activity on plate tests unless cyclic AMP was included on the plate. Figure 5 shows the electrophoretic analysis of polypeptides synthesized by strain DG74 transformed by pBGP120, pBGP123, or pBGP124. Again, the basal levels of all the plasmid-derived β-galactosidase polypeptide bands are very low but inducible. These data show that the *lac* promoter and operator in pBGP120 and hybrid derivatives are subject to regulation by the same effectors that control the chromosomal *lac* elements: the *lac* repressor and cyclic AMP.

DISCUSSION

The plasmid vehicle pBGP120 carries the *lac* operator, promoter, and most of the β-galactosidase structural gene (17 and below). The β-galactosidase polypeptide encoded by the plasmid is shown to have low enzymatic activity and altered electrophoretic mobility. These data suggest the presence of an *Eco*RI site within the β-galactosidase gene of *E. coli*. If this is so, the normal translational termination signal would have been deleted from the β-galactosidase gene in the construction of pBGP120. Consequently, transcription and translation of *lacZ* should read through the *Eco*RI site and express sequences distal to *lacZ*. The size of the β-galactosidase monomer produced from the plasmid would then depend on the position of the first in-phase nonsense codon in the DNA distal to *lacZ*. This prediction was tested by the construction and analysis of hybrid plasmids, each of which car-

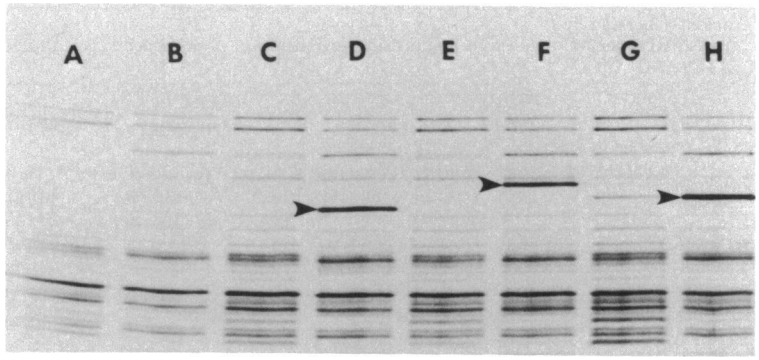

FIG. 4. *SDS-gel electrophoresis demonstrating IPTG induction of plasmid-encoded β-galactosidase monomer and related polypeptides in transformants of strain M96. Cells grown in glucose minimal media were labeled with or without a prior induction with cyclic AMP and IPTG. The samples (A) M96, (B) induced M96, (C) M96 (pBGP120), (D) induced M96 (pBGP120), (E) M96 (pBGP123), (F) induced M96 (pBGP123), (G) M96 (pBGP124), and (H) induced M96 (pBGP124) were subjected to electrophoresis on a 7.5% acrylamide gel. A portion of an autoradiogram of the resulting separation is shown.*

ries a different DNA sequence distal to *lacZ*. Each was shown to produce a β-galactosidase-related polypeptide of a different size. Thus, there is an *Eco*RI site within *lacZ*.

The known amino acid sequence of the β-galactosidase polypeptide (2) indicates five possible positions at which an *Eco*RI site could occur (listed in Table 3). The following argument shows that the *Eco*RI site must be near the distal end of *lacZ*. The sizes of the different β-galactosidase-related polypeptides produced from pBGP120 and hybrid plasmids carrying inserted DNA sequences were all as large or larger than the wild-type β-galactosidase monomer. Thus, at a minimum, the number of amino acids resulting from readthrough translation must be sufficient to compensate for the deletion of the distal end of the coding sequence. If the DNA expressed by readthrough translation is normally not translated and we assume a random DNA sequence, we can calculate the probability of adding the required number of amino acids. If the *Eco*RI site is 16 amino acids from the COOH terminus (Table 3), the probability of producing a product as large or larger than the normal subunit is 0.46 (determined by taking the probability of adding an amino acid, 61/64, to the 16th power). On the other hand, if the *Eco*RI site were 478 amino acids from the COOH terminus (Table 3), this probability would be less than 1.1×10^{-10} $[(61/64)^{478}]$. Thus, we conclude that the *Eco*RI site very likely codes for

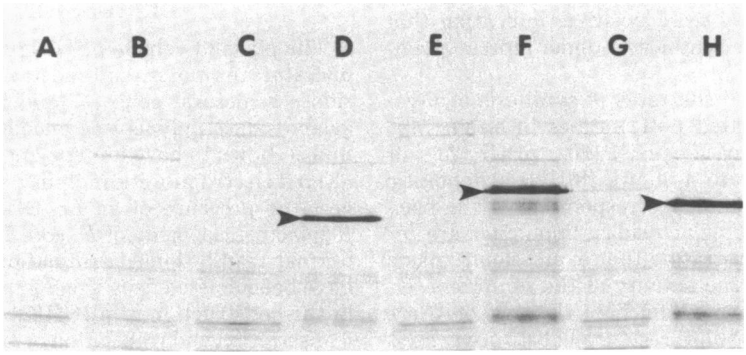

FIG. 5. *SDS-gel electrophoresis demonstrating induction of plasmid β-galactosidase monomer and related polypeptides in transformants of an adenyl cyclase-deficient strain. Cells grown in glucose minimal media were labeled with or without prior exposure to cyclic AMP and IPTG. The samples (A) DG74, (B) induced DG74, (C) DG74(pBGP120), (D) induced DG74(pBGP120), (E) DG74(pBGP123), (F) induced DG74(pBGP123), (G) DG74(pBGP124), and (H) induced DG74(pBGP124), were subjected to electrophoresis on a 7.5% acrylamide gel. A portion of the autoradiogram of the resulting separation is shown.*

TABLE 3. *Phase designations, codon assignments and potential EcoRI sites in β-galactosidase*

Phase designation[a]	Codons generated by the *Eco*RI sequence XXG↓AAUUCXX[b]	Possible amino acid assignments	Sequence	Position[c]	Amino acids deleted[d]
+1	G↓AA	Glu	Glu-Phe	107–171	850
	UUC	Phe	Glu-Phe	1004–1005	16
0	XXG↓	All amino acids except: Asn, Asp, Cys, His, Ile, Phe, Tyr	Arg-Asn-Ser	37–39	982
			Val-Asn-Ser	146–148	873
			Gly-Asn-Ser	541–543	478
	AAU	Asn			
	UCX	Ser			
−1	XG↓A	Arg, Gly, (Opal)	None		
	AUU	Ile			
	CXX	Leu, Pro, His, Arg, Gln			

[a] See text for description.

[b] Arrow indicates position of *Eco*RI endonuclease cleavage. Positions at which the nucleotide cannot be designated are indicated with an X.

[c] From sequence data of Fowler and Zabin (2). Numbers indicate specific amino acid positions in β-galactosidase.

[d] Number of amino acids deleted if the *Eco*RI site is at the designated position.

Glu-Phe at positions 1004 and 1005 of the β-galactosidase monomer. *Eco*RI cleavage of *lacZ* removed the coding sequence for the COOH-terminal 18 amino acids. However, the Glu-Phe codons will be regenerated by ligation to another *Eco*RI-cleaved molecule.

The conclusion that the *Eco*RI site codes for the amino acids Glu-Phe defines the position of the restriction site in relation to the translational phase. However, no systematic way of designating this "phase" has been established. Therefore, we have defined the phase of a breakpoint as follows: if the antisense strand of DNA (whose sequence corresponds to the mRNA sequence) is broken between two codons, the phase is 0; if the break occurs between positions 1 and 2 of a codon, the phase is +1; and if the break occurs between positions 2 and 3 of a codon, the phase is −1 (see Table 3).

By using this system, the two DNA termini generated by a single cleavage of a coding sequence will be designated by the same phase. To indicate the polarity of transcription, the end of a DNA fragment can be further designated as terminating with a free 3′ or 5′ end in the antisense strand. Thus, the end of the β-galactosidase coding sequence in pBGP120 is designated as 3′+1. Clearly, ligation of 3′+1 and 5′+1 ends is necessary to generate a hybrid DNA sequence capable of regenerating the original reading frame. Because readthrough translation is a potential mechanism of generating expression of foreign DNA, this nomenclature should prove useful.

The last 16 amino acids of the β-galactosidase polypeptide are essential for activity, because a strain carrying a complete *lacZ* deletion could not be transformed to Lac$^+$ by pBGP120. Mutations in the terminal third of *lacZ* produce an enzymatically inactive peptide that can be activated by a product, omega peptide, produced by mutations that leave the COOH terminal third of *lacZ* intact (20). Because strain DG73 probably carries a small deletion in the proximal one-third of *lacZ*, it would be capable of omega complementation.

The data presented here show that the truncated *lacZ* carried by pBGP120 is under the control of the *lac* operator. Its expression is cyclic AMP-dependent and can be repressed by the *lac* repressor. The induced level of expression of the plasmid *lacZ* is about sevenfold higher than the induced level of expression of chromosomal *lacZ*, suggesting that expression increases in proportion to the gene copy number. Expression of sequences distal to *lacZ* is also under the control of the *lac* operator.

Several possible selection and screening procedures should enhance the value of pBGP120 as a cloning vehicle. One possible approach assumes that insertion of noncoding DNA distal to *lacZ* will result in the addition of relatively few amino acids to β-galactosidase and, therefore, will be unlikely to alter the enzyme activity. In contrast, the in-phase insertion of coding DNA will usually result in a large addition of β-galactosidase and will be likely to sterically interfere with the complex quaternary structure required for omega complemented activity. As predicted by these assumptions, the insertion of the noncoding *Xenopus* rDNA into pBGP120 did not interfere with the capacity of the plasmid to confer a Lac$^+$ phenotype to strain DG73. In contrast, Lac(Ts) transformant, clone 8, and a Lac$^-$ transformant, clone 12, produced very large *lacZ*-related polypeptides. If the inserted DNAs were random sequences, the probability of adding 48,000 daltons (clone 12) or 28,000 daltons (clone 8) to β-galactosidase would be less than 5×10^{-9} and 1.4×10^{-5}, respectively. These calculations suggest that the sequences inserted in pBGP120 in clones 8 and 12 are in-phase coding DNA. The successful isolation of in-phase insertions adequately demonstrates the power of this selection procedure even though the cloned DNA fragments are from an uncharacterized source.

In addition, the ability to regulate the expression of distal coding sequences in pBGP120 using *cya* or *lacI*(Sq) backgrounds should allow for the cloning of DNA sequences lethal to the cell if expressed constitutively.

Finally, it is worth noting that in all cases observed thus far, the fused polypeptides generated by readthrough translation in pBGP120 not only have a high rate of synthesis but accumulate to high levels and, therefore, must be quite stable in vivo. By contrast, it has been generally observed that gross alterations in protein structure resulting from mutation or chemical alteration lead to highly unstable polypeptides (22). The basis for the stability of the plasmid-encoded fused polypeptides is unknown.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation grant GMS75-19418 to B. J. McCarthy and Public Health Service grant GM17239 from the National Institute for General Medical Sciences to the late G. M. Tomkins.

We thank K. Y. Wong for technical assistance, A. Riggs, D. Zipser, and D. Morse for providing strains, L. DeGennaro for *Eco*RI endonuclease, G. Bell for CD4 DNA, and I. Zabin, R. Tait, and Mark Alper for communicating results before publication. We thank B. Alberts, B. J. McCarthy, and S. N. Cohen for critical comments on the manuscript.

LITERATURE CITED

1. **Clewell, D. B.** 1972. Nature of Col E_1 plasmid replication in *Escherichia coli* in the presence of chloramphenicol. J. Bacteriol. **110**:667-676.

2. **Fowler, A. V., and I. Zabin.** 1977. The amino acid sequence of β-galactosidase of *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **74**:1507-1510.
3. **Gilbert, W., and B. Müller-Hill.** 1970. The lactose repressor, p. 93-109. *In* J. R. Beckwith and D. Zipser (ed.), The lactose operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
4. **Helling, R. B., H. M. Goodman, and H. W. Boyer.** 1974. Analysis of endonuclease R·*Eco*RI fragments of DNA from lambdoid bacteriophages and other viruses by agarose-gel electrophoresis. J. Virol. **14**:1235-1244.
5. **Hershfield, V., H. W. Boyer, C. Yanofsky, M. A. Lovett, and D. R. Helinski.** 1974. Plasmid ColE1 as a molecular cloning vehicle for cloning and amplification of DNA. Proc. Natl. Acad. Sci. U.S.A. **71**:3455-3459.
6. **Hudson, B., W. B. Upholt, J. Devinny, and J. Vinograd.** 1969. The use of an ethidium analogue in the dye-bouyant density procedure for the isolation of closed circular DNA: the variation of the superhelix density on mitochondrial DNA. Proc. Natl. Acad. Sci. U.S.A. **62**:813-820.
7. **Ivarie, R. D., D. H. Gelfand, P. P. Jones, P. Z. O'Farrell, B. A. Polisky, R. A. Steinberg, and P. H. O'Farrell.** 1977. Biological applications of two-dimensional gel electrophoresis, p. 369-384. *In* B. J. Radola and D. Graesslin (ed.), Electrofocusing and isotachophoresis. Walter de Gruyter and Co., New York.
8. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680-685.
9. **Laemmli, U. K., and M. Favre.** 1973. Maturation of the head of bacteriophage T4. I. DNA packaging events. J. Mol. Biol. **80**:575-599.
10. **Meagher, R. B., R. C. Tait, M. Betlach, and H. W. Boyer.** 1977. Protein expression in *E. coli* minicells by recombinant plasmids. Cell **10**:521-536.
11. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
12. **Morrow, J. F., S. N. Cohen, A. C. Y. Chang, H. W. Boyer, H. M. Goodman and R. B. Helling.** 1974. Replication and transcription of eukaryotic DNA in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **71**:1743-1747.
13. **National Institutes of Health.** 1976. Guidelines for research involving recombinant DNA molecules. Fed. Reg. **41**:27902-27943.
14. **O'Farrell, P. H.** 1975. High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. **250**:4007-4021.
15. **O'Farrell, P. H., and P. Z. O'Farrell.** 1976. Two-dimensional polyacrylamide gel electrophoretic fractionation. Methods Cell Biol. **16**:407-420.
16. **Perlman, R. L., B. DeCrombrugghe, and I. Pastan.** 1969. Cyclic AMP regulates catabolite and transient repression in *E. coli*. Nature (London) **223**:810-812.
17. **Polisky, B., R. J. Bishop, and D. H. Gelfand.** 1976. A plasmid cloning vehicle allowing regulated expression of eukaryotic DNA in bacteria. Proc. Natl. Acad. Sci. U.S.A. **73**:3900-3904.
18. **Shapiro, J., L. Machattie, L. Eron, G. Ihler, K. Ippen, and J. Beckwith.** 1969. Isolation of pure *lac* operon DNA. Nature (London) **224**:768-774.
19. **So, M., R. Gill, and S. Falkow.** 1975. The generation of a ColE1-Ap^R cloning vehicle which allows detection of inserted DNA. Mol. Gen. Genet. **142**:239-249.
20. **Ullman, A., and D. Perrin.** 1970. Complementation in β-galactosidase, p. 143-172. *In* J. R. Beckwith and D. Zipser (ed.), The lactose operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
21. **Vogel, H. J., and D. M. Bonner.** 1956. Acetylornithinease of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. **218**:97-106.
22. **Zipser, D., and P. Bhavsar.** 1976. Missense mutations in the *lacZ* gene that result in degradation of β-galactosidase structural protein. J. Bacteriol. **127**:1538-1542.