JOURNAL OF BACTERIOLOGY, June 1987, p. 2327–2335
0021-9193/87/062327-09$02.00/0

Vol. 169, No. 6

Gene Expression in *Zymomonas mobilis*: Promoter Structure and Identification of Membrane Anchor Sequences Forming Functional LacZ′ Fusion Proteins†

T. CONWAY, Y. A. OSMAN, AND L. O. INGRAM*

Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida 32611

Received 26 August 1986/Accepted 26 February 1987

We have described a procedure for the isolation of *lacZ′* fusion genes which contain anchor sequences conferring membrane association. This method was used to isolate fragments of DNA from *Zymomonas mobilis* which contain promoter activity and amino-terminal sequences. The sequences and transcriptional initiation sites of three of these were compared. Both *Escherichia coli* and *Z. mobilis* recognized similar regions of DNA for transcriptional initiation. Five to eight consecutive hydrophobic amino acids in the amino terminus served to anchor these hybrid proteins to the membrane in both *E. coli* and *Z. mobilis*. General features observed in the *Z. mobilis* fragments included partial sequence homology with the −35 region sequence of *E. coli*, repetitive and palindromic A+T-rich regions preceding and adjoining the −10 region, a sequence resembling the consensus sequence of *E. coli* in the −10 region, and a potential ribosomal-binding site (AGGA) 8 to 12 bases upstream from an in-frame start codon. The level of expression of fusion proteins was generally higher in *E. coli* than in *Z. mobilis*. This higher level of expression in *E. coli* may result from multiple sites of transcriptional initiation and higher plasmid copy number.

Zymomonas mobilis is an unusual gram-negative bacterium of uncertain taxonomic position (23). This organism, alone, is obligately fermentative and utilizes an Entner-Doudoroff pathway for glycolysis (39). It exhibits an ethanol tolerance equivalent to that of yeasts and is capable of rapid ethanol production (29), making it potentially useful for the commercial production of fuel ethanol.

Little is known about promoter structure and gene expression in this organism. Previous studies have shown that the tetracycline gene contained in broad-host-range plasmids such as RP4, pGC91.14, pRK290, and pCVD305 are expressed at sufficient levels to allow selection (7, 37, 38). Other studies have shown that the *lacZ* gene on transposon Tn*951* is expressed poorly even in the presence of inducer (7). Recent studies in our laboratory have indicated that the chloramphenicol resistance gene is expressed in *Z. mobilis* (T. Conway and L. O. Ingram, Abstr. Annu. Meet. Am. Soc. Microbiol. 1986, H62, p. 137) under the control of the *Escherichia coli* consensus promoter (31). No studies have been reported concerning the sequence requirements for membrane localization or protein export in *Z. mobilis*.

In this study, we investigated promoter structure in *Z. mobilis* by isolating *lacZ* gene fusions in which the promoter, translational start sequence, and N-terminal peptide sequence are provided by fragments of *Z. mobilis* DNA. High levels of β-galactosidase have been reported previously to confer a lactose-positive phenotype in the absence of functional lactose permease (33), and this feature was used to bias our selection for hybrids which exhibited high levels of β-galactosidase activity, i.e., *Z. mobilis* fragments with strong promoter activity. Over half of the hybrids selected in this way were found to be associated with the cell membrane.

* Corresponding author.
† Florida Agricultural Experiment Station publication no. 7469.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The bacterial strains and plasmids used in this study are summarized in Table 1. *Z. mobilis* CP4 and derivatives containing plasmids were grown at 30°C in complex medium as previously described (5). Strains of *E. coli* were grown at 37°C in Luria broth (18) containing no carbohydrate, except when otherwise indicated. Strains of *E. coli* were screened for lactose utilization on MacConkey agar plates containing 10 g of lactose per liter or on Luria agar plates containing 20 mg of 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) per liter. Strains of *Z. mobilis* were screened on complex medium containing 20 mg of X-Gal per liter. Chromosomal β-galactosidase was induced in *E. coli* HB101 by growth in the presence of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). A spontaneous *lacZY* derivative of *E. coli* S17-1 was isolated following enrichment during growth with *o*-nitrophenyl-β-D-thiogalactopyranoside, as described by Miller (22), and was designated strain TC4.

Transformation and conjugation. *E. coli* strains were transformed by the $CaCl_2$ procedure of Mandel and Higa (19) and were selected for resistance to chloramphenicol (40 mg/liter). Plasmid shuttle vectors carried in *E. coli* TC4 were conjugated into *Z. mobilis* by using the membrane filter procedure described by Simon et al. (36). After incubation, cells were washed from filters and plated on complex medium containing nalidixic acid (20 mg/liter) to inhibit the growth of *E. coli* and chloramphenicol (120 mg/liter) to select plasmid-containing strains of *Z. mobilis*. *Z. mobilis* exhibits natural resistance to nalidixic acid (23).

β-galactosidase assays. Assays for β-galactosidase were performed with *o*-nitrophenyl-β-D-galactopyranoside (ONPG) by the method of Miller (22). Activities are reported in Miller units. Samples of spent medium were obtained after sedimenting mid-log-phase cells by centrifugation in an

TABLE 1. Plasmids and strains used in this study

Organism	Relevant genotype and description	Source or reference
Z. mobilis CP4	Prototroph	A. Ben-Bassat; 6
E. coli		
HB101	*lacYI recA rpsL*	D. Duckworth[a]; 19
S17-1	Mobilizing strain	R. Simon; 35
TC4	*lacZY* derivative of S17-1	This laboratory
Plasmids		
RSF1010	Broad-host-range replicon mob site	Roy Curtiss III[b]; 2, 13
pXJ002	Consensus promoter-CAT	ATCC[c]; 26
pSUP205	mob site of RP4 CAT (Cm^r)	R. Simon; 35
pORF2	truncated *lacZ'* colE1 origin	G. Weinstock; 39
pLOI245	Fusion vector	This paper
pLOI193	pLOI245 with Tet^r replacing *lacZ*	9
pLOI301	*Z. mobilis-lacZ'* fusion derivative of pLOI245	This paper
pLOI302	*Z. mobilis-lacZ'* fusion derivative of pLOI245	This paper
pLOI303	*Z. mobilis-lacZ'* fusion derivative of pLOI245	This paper

[a] University of Florida, Gainesville.
[b] Washington University, St. Louis, Mo.
[c] ATCC: American Type Culture Collection, Rockville, Md.

Eppendorf centrifuge for 1 min. Cells were washed, suspended in Z buffer (22), and assayed without permeabilization as a measure of surface-accessible activity in *lacY* strains of *E. coli* (grown with 2 g of glucose per liter to repress chromosomal β-galactosidase) and in recombinant strains of *Z. mobilis*. Cells were permeabilized by treatment with chloroform and sodium dodecyl sulfate to determine total β-galactosidase activity (22).

Localization of β-galactosidase activity. To isolate cytoplasmic and membrane fractions, 1-ml samples of mid-log-phase cells were harvested by centrifugation (10,000 × *g*, 1 min), suspended in an equal volume of 100 mM Tris hydrochloride (pH 8.0) containing 10 μg of lysozyme per ml, and incubated for 1 h at 30°C. After incubation, DNase I and $MgCl_2$ were added (final concentrations, 5 μg/ml and 10 mM, respectively), and cells were disrupted during seven freeze-thaw cycles (−60 to 30°C), frozen overnight at −20°C, and thawed slowly on ice. Such extensive freeze-thaw cycling was in excess of that required for *E. coli* but was essential to disrupt *Z. mobilis*. The membrane fractions were collected by centrifugation in an Eppendorf microcentrifuge for 5 min at 5°C. The supernatant represented the cytoplasmic fraction and contained 70% of the total cell protein. The membrane pellet was thoroughly suspended in an equal volume of 100 mM Tris hydrochloride (pH 8.0)–10 mM $MgCl_2$ buffer and is referred to as the membrane fraction. For comparison, total activity was measured without centrifugation in parallel cell lysates. Periplasmic shock proteins were obtained by the method of Willis et al. (42), and the β-galactosidase activity in the shock fluid was measured. Membrane localization was also confirmed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis of cytoplasmic and membrane fractions, essentially as described by Weinstock et al. (40).

The localization of β-galactosidase fusion proteins was further examined by treatment of *E. coli* spheroplasts with proteinase K by a modification of the procedure of Randall (27). Cells were harvested in mid-log phase and used immediately. All incubations were at 30°C, and phenylmethylsulfonyl fluoride was omitted. Samples were taken at 5-min intervals after the addition of proteinase K and were immediately assayed for remaining β-galactosidase activity without permeabilization.

DNA isolation and constructions. Small-scale isolations of plasmid DNA from *E. coli* strains were prepared as described by Silhavy et al. (35). Large-scale plasmid isolations were prepared by the alkaline sodium dodecyl sulfate procedure described by Maniatis et al. (20), and the plasmids were purified twice on cesium chloride-ethidium bromide gradients. Genomic DNA from *Z. mobilis* was isolated as described by Byung et al. (5). Restriction endonucleases and T4 DNA ligase were used as recommended by the manufacturers. Restriction fragments were analyzed by agarose gel electrophoresis (20).

DNA sequence analysis. The base sequences of three fragments of *Z. mobilis* DNA were determined by the dideoxy sequencing method (32) using collapsed plasmids (8, 41) as described previously (9). The M13 sequencing primer used (5'-CAGCACTGACCCTTTTG-3') is homologous to a region of the *lacZ* gene, 19 bases downstream from the point of fusion (*Bam*HI site) with the inserted fragments of *Z. mobilis* DNA.

Analyses of transcript 5' termini. The 5' termini of the transcripts from *Z. mobilis* promoter-*lacZ'* fusion genes were determined by primer extension analysis (25) and S1 nuclease mapping (11). Primer extension experiments were carried out as previously described (9). Total RNA was isolated from mid-log-phase cells as described previously (4, 9).

S1 nuclease mapping experiments were performed by the procedure of Maniatis et al. (20). The DNA used for hybridization to RNA was located on a *Pvu*II-*Hin*dIII restriction fragment for each of the three promoters examined. The *Pvu*II site is in the native portion of the *lacZ'* hybrid genes. The *Hin*dIII site and *Cla*I site are in the vector, upstream of the inserted *Z. mobilis* fragments. *Pvu*II-*Cla*I DNA fragments were 5' end labeled (20), digested with *Hin*dIII to remove label from the *Cla*I end, and purified by agarose gel electrophoresis.

Hydropathy indices of N termini of β-galactosidase fusion proteins. Hydropathy indices were plotted by using the method of Kyte and Doolittle (17). These were computed by using the DNA and protein sequence management program of Pustell and Kafatos (26) with an averaging window of 9 amino acids.

Plasmid copy number. Cells (50-ml cultures of *E. coli* and 100-ml cultures of *Z. mobilis*) were labeled by growth in Luria broth containing 0.5 to 1.0 μCi of $^{32}PO_4^{3-}$ per ml. Cells were harvested, washed twice, and suspended in 10 mM Tris hydrochloride buffer containing 1 mM EDTA (pH 8.0). Suspensions were added dropwise with agitation to 30 ml of 4% sodium dodecyl sulfate containing 10 mM Tris hydrochloride and 1 mM EDTA (pH 12.4). Cells were lysed by incubation at 37°C for 20 min with gentle agitation, neutralized with HCl to pH 8.0, and digested at 37°C for 5 h with DNase-free RNase (10 μg/ml) and then for 16 h with protease K (200 μg/ml). Plasmid and chromosomal DNA were separated on 0.8% agarose gels and counted with a model LS8000 scintillation counter (Beckman Instruments, Inc., Fullerton, Calif.).

Enzymes, chemicals, and reagents. Restriction enzymes and sequence management programs were purchased from International Biotechnologies, Inc., New Haven, Conn. Avian myeloblastosis virus reverse transcriptase was pur-

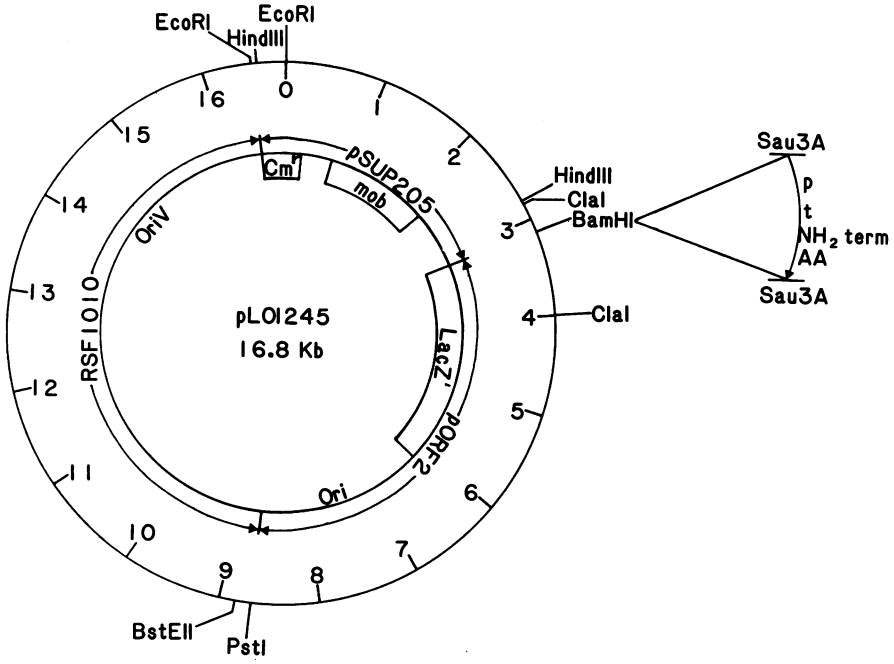

FIG. 1. Physical map of the shuttle vector pLOI245 used for construction of β-galactosidase gene fusions. Portions of four plasmids were used in the construction of pLOI245. Three of these are shown, with arrows delineating the points of fusion. A small *Eco*RI fragment from a fourth plasmid, pXJ002, provided a consensus promoter and N-terminal segment of the chloramphenicol resistance gene (not marked). For construction of β-galactosidase fusions, pLOI245 was opened at the unique *Bam*HI site, and *Sau*3A fragments of *Z. mobilis* genomic DNA (containing a promoter, translational start, and N-terminal amino acids) were ligated into this site.

chased from Boehringer Mannheim Biochemicals, Indianapolis, Ind. The M13 dideoxy sequencing kit, Klenow fragment of DNA polymerase I, and S1 nuclease were obtained from Bethesda Research Laboratories, Inc., Gaithersburg, Md. M13 sequencing primer (17-mer) was obtained from New England BioLabs, Inc., Beverly, Mass. Radioactive compounds were purchased from New England Nuclear Corp., Boston, Mass.

RESULTS

Construction of vector. We have constructed a broad-host-range vector, pLOI245, for the selection of *lacZ'* fusion genes (Fig. 1). The broad host replication functions for this vector are contained on a 7.8-kilobase (kb), *Pst*I-*Eco*RI fragment from the plasmid RSF1010. Chloramphenicol serves as the selectable marker, and this gene was constructed from a small *Eco*RI-*Eco*RI fragment from pXJ002 containing the consensus promoter and N-terminal end of the CAT gene and a 3-kb *Eco*RI-*Bam*HI fragment from pSUP205. This pSUP205 fragment also provided a site for plasmid mobilization. The promoterless, N-truncated β-galactosidase gene (lacking the first eight amino acids) was derived from pORF2 as a 5.6-kb *Bam*HI-*Pst*I fragment. This fragment from pORF2 also included a colE1 replicon. The unique *Bam*HI site on this vector was used to clone *Sau*3A fragments of *Z. mobilis* DNA which completed the β-galactosidase gene.

Selection of fusion genes conferring a Lac$^+$ phenotype in the absence of a functional lactose permease. Although native β-galactosidase is a cytoplasmic protein, recent studies with hybrid genes in *Saccharomyces cerevisiae* (33) and in *Streptomyces lividans* (T. Eckhardt, M. Rosenberg, and L. Fare, Abstr. Am. Soc. Microbiol. First Annu. Conf. Biotechnol. 1986, abstr. no. 218, p. 26) have shown that sequences specifying membrane localization and export can be added to the amino-terminal end with the retention of enzyme activity in some systems. We have ligated *Sau*3A partial-digestion products of *Z. mobilis* DNA (predominantly 2 kb and smaller) into the *Bam*HI site of pLOI245. These were transformed into *E. coli* TC4 and *E. coli* HB101, which lack a functional lactose permease gene. Clones were selected which exhibited a strong lactose-positive phenotype (approximately 0.3% of the transformants) on lactose-MacConkey agar plates containing chloramphenicol. These clones contained fusion genes, which circumvented the need for lactose permease in *E. coli*, as evidenced by passage of this trait to progeny during subsequent transformations with isolated plasmid DNA.

Of the 300 clones isolated, 17 were selected for further study. The *Sau*3A inserts in these clones ranged in size from 7.6 to 0.2 kb. On the basis of size and restriction sites, each of these clones appeared unique. These plasmids were subsequently transferred into *Z. mobilis* CP4 by conjugation, with *E. coli* TC4 as the mobilizing donor. The presence of functional β-galactosidase in *Z. mobilis* was confirmed on plates containing X-Gal.

Expression of fusion genes in *E. coli* and *Z. mobilis*. The total β-galactosidase activities exhibited by clones of *E. coli* and *Z. mobilis* containing the fusion genes were measured after permeabilization with chloroform and detergent (Table 2). The activities present in *E. coli* were unusually high, 5- to 30-fold higher than the levels observed in a fully induced *E. coli* HB101, and ranged from 3,049 to 23,131 Miller units. In *Z. mobilis*, most gene fusions exhibited less than half of the activity observed in *E. coli*, ranging from 633 to 6,404 Miller units. In general, the comparative level of expression among the different fusions followed the same trends in both orga-

TABLE 2. β-Galactosidase activity of permeabilized cells, intact cells, and spent medium[a]

Clone		Activity in *E. coli*[b]			Activity in *Z. mobilis*		
		Permeabilized cells	Intact cells	Spent medium	Permeabilized cells	Intact cells	Spent medium
	A	10,937	217	128	5,186	25	4
(pLOI301)	B	9,439	262	117	4,992	40	5
	C	11,311	285	49	4,453	28	1
	D	12,369	987	35	2,304	11	0
	E	4,137	190	40	3,280	19	1
	F	8,808	231	20	5,525	55	0
(pLOI302)	G	8,661	529	6	1,226	5	0
	H	15,373	361	38	964	6	0
	I	11,988	274	63	5,586	41	4
(pLOI303)	J	23,131	349	59	5,141	37	2
	K	14,663	426	28	4,816	37	2
	L	6,024	277	19	872	2	0
	M	7,458	603	32	1,076	3	0
	N	3,049	303	11	1,958	0	0
	O	4,398	304	2	552	44	5
	P	3,097	112	5	6,404	56	1
	Q	5,652	227	ND[c]	633	3	0
E. coli HB101[d]		786	53	1			
Z. mobilis CP4					0		

[a] Activity is expressed as Miller units.
[b] *E. coli* was grown in Luria broth containing 2 g of glucose per liter to repress the synthesis of native β-galactosidase.
[c] ND, Not determined.
[d] *E. coli* HB101 was grown in the presence of IPTG without added glucose to induce β-galactosidase.

nisms. However, several exceptions should be noted. Clone P exhibited higher activity in *Z. mobilis* than in *E. coli*. Clone E was expressed in *Z. mobilis* at over 75% of the level found in *E. coli*. Clone H was expressed in *Z. mobilis* at less than 10% of the level found in *E. coli*.

Copy number was examined for a related plasmid, pLOI193, which contains a tetracycline gene instead of a truncated *lacZ* (9). This plasmid was present in *Z. mobilis* CP4 at 11 to 30 copies per genome equivalent. A much higher copy number was observed in *E. coli* HB101, over 100 copies per genome equivalent. It is likely that the higher copy number is in part responsible for the higher β-galactosidase activities observed in *E. coli*.

The entry of ONPG into native (nonpermeabilized) cells serves to limit its rate of hydrolysis by β-galactosidase in *E. coli* strains which contain functional lactose operons. This serves as the basis for a sensitive method to determine lactose permease activity and ONPG uptake by other mechanisms, termed cryptic transport systems (12). Strain HB101 contains a functional β-galactosidase gene but does not contain a functional permease gene. The level of β-galactosidase activity observed in this strain lacking fusion genes, 53 Miller units, reflects the rate of entry of ONPG mediated by the cryptic transport systems. This level of activity is not sufficient to produce the red, lactose-positive phenotype on lactose-MacConkey plates. The hybrid β-galactosidase genes which were selected as having a lactose-positive phenotype in strain HB101 all expressed higher levels of activity in the absence of permeabilization (Table 2). These ranged from 2- to 10-fold that of the control strain lacking plasmid. However, these activities represented less than 10% of the total activity, indicating that added ONPG was not readily accessible to the bulk of the hybrid β-galactosidase protein. Very low levels of β-galactosidase activity were detected in strains of *Z. mobilis* containing the fusion genes in the absence of cell permeabilization, usually less than the levels observed as cryptic uptake in *E. coli* HB101. Recombinant strains of *Z. mobilis* did not grow on complex medium containing lactose as the fermentable sugar, presumably owing to the lack of an efficient uptake system for lactose or its analog, ONPG.

A low percentage of the β-galactosidase activity from the hybrid genes was observed in the culture medium, particularly with *E. coli*. This may be due to cell lysis during growth or to cell breakage during centrifugation. Although the activities were low in most cases, it is possible that they contributed to the lactose-positive phenotype observed with recombinants of *E. coli* during growth on lactose-MacConkey agar by providing an exogenous supply of glucose and galactose.

Cellular localization of hybrid β-galactosidase proteins. We have examined the location of the hybrid β-galactosidase proteins in both *Z. mobilis* and *E. coli* (Table 3). None of the clones contained appreciable levels of activity in the periplasmic protein fraction (less than 0.8%; data not shown). Strain HB101 was examined as a control. In this strain, 1% of the total activity was recovered in the periplasmic fraction. The activities of seven of the hybrid genes were located in the cytoplasmic fraction, and the activities of four were predominantly membrane bound. The activities of six of the hybrid genes were divided, with major portions in both the cytoplasmic and membrane fractions. Representative fusion genes were selected from each group for further study. These constructions were designated pLOI301 (clone B; cytoplasmic activity), pLOI302 (clone G; membrane-bound activity), and pLOI303 (clone J; activity divided between cytoplasmic and membrane fractions).

The location of the fusion proteins in *E. coli* was further confirmed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis of cytoplasmic and membrane fractions (Fig. 2). The fusion protein from pLOI301 was predominantly in the cytoplasmic fraction, that of pLOI302 was predomi-

TABLE 3. Intracellular location of β-galactosidase activity[a]

Clone		Activity in *E. coli*[b]		Activity in *Z. mobilis*		Membrane localized[c]
		Membrane bound	Cytoplasmic	Membrane bound	Cytoplasmic	
	A	3.1	96.9	5.6	94.4	
(pLOI301)	B	4.0	96.0	6.3	93.7	
	C	4.1	95.9	5.9	94.1	
	D	54.0	46.0	78.1	21.9	*
	E	3.0	97.0	8.3	91.7	
	F	32.7	67.3	45.7	54.3	*
(pLOI302)	G	88.5	11.5	94.1	5.9	**
	H	18.5	81.5	26.8	73.2	*
	I	2.0	98.0	11.0	89.0	
(pLOI303)	J	39.7	60.3	59.8	40.2	*
	K	60.1	39.9	52.0	48.0	*
	L	77.6	22.4	93.1	6.9	**
	M	14.4	85.6	11.8	88.2	
	N	15.3	84.7	16.3	83.7	
	O	94.0	6.0	92.3	7.7	**
	P	33.1	66.9	61.5	38.5	*
	Q	87.1	12.9	83.9	16.1	**
E. coli HB101[d]		3.7	96.3	—	—	

[a] β-Galactosidase activity was measured in Miller units and is expressed as a percentage of total activity.
[b] Cells were grown in Luria broth containing 2 g of glucose per liter to repress the synthesis of native β-galactosidase.
[c] **, high degree of membrane association; *, intermediate level of membrane association.
[d] *E. coli* HB101 were grown in the presence of IPTG with glucose.

nantly in the membrane fraction, and that of pLOI303 was divided between cytoplasmic and membrane fractions.

Washed, nonpermeabilized cells of *E. coli* containing the representative fusion genes each exhibited appreciable β-galactosidase activity, although no lactose permease was present (Table 2). To determine whether this activity represented a portion of the hybrid β-galactosidase which had been exported to the outer membrane or the outer surface of the plasma membrane, we examined the effects of protease K treatment on stabilized spheroplasts. Proteinase K did not reduce the activity of β-galactosidase in osmotically stabilized spheroplasts during 40 min of incubation at 30°C, indicating that the activity observed in nonpermeabilized cells was not due to exposed hybrid enzyme. When spheroplasts were treated with chloroform, β-galactosidase activity was rapidly lost, demonstrating that the hybrid proteins were readily digested by proteinase K after disruption of the plasma membrane. These results indicate that the activity observed in nonpermeabilized cells is not due to the localization of functional hybrid β-galactosidase on the outer surface of the plasma membrane, on the outer membrane, or in the periplasmic space.

DNA sequences of three clones. The DNA base sequences for the amino-terminal segment (derived from *Z. mobilis*) of the three representative hybrid genes were determined by collapsed-plasmid sequencing (Fig. 3). The positions of the translational start signals were inferred on the basis of transcriptional mapping (discussed below). Potential ribosomal-binding sites were identified by inspection and found to be immediately upstream from the first in-frame start codons in pLOI302 and pLOI303. The positions of translational initiation in pLOI301 and pLOI303 are ambiguous, with two possible in-frame start codons, an ATG and a GTG. In pLOI301, the first of these (Fig. 3A, base position [bp] 18) is not preceded by a sequence resembling a ribosome-binding site (16). The second methionine occurs at bp 75 and is preceded by the sequence AGGA; it represents the probable site for translational initiation. In pLOI303, only the first start codon (ATG) is appropriately spaced from a ribosomal-binding sequence and is the probable site of translational initiation.

The amino terminus of the β-galactosidase fusion protein encoded by *Z. mobilis* DNA in pLOI301 (Fig. 3A) consists mainly of hydrophilic amino acids, whether translation is initiated from the first or the second in-frame methionine codon. The longest consecutive stretch of hydrophobic amino acids in this fusion is three. The β-galactosidase fusion protein from pLOI301 was found to remain freely soluble in the cytoplasm. In contrast, the fusion protein from pLOI302 (Fig. 3B) consisted primarily of hydrophobic amino acids and was associated with the cell membrane when produced in either *E. coli* or *Z. mobilis* (88 and 94%, respectively). The amino terminus of this fusion protein contained a stretch of eight hydrophobic amino acids. The amino terminus of the fusion protein from pLOI303 (Fig. 3C) consisted of a mixture of hydrophilic and hydrophobic amino acids and contained a stretch of five consecutive hydrophobic amino acids. Approximately 50% of this fusion protein was associated with the cell membrane.

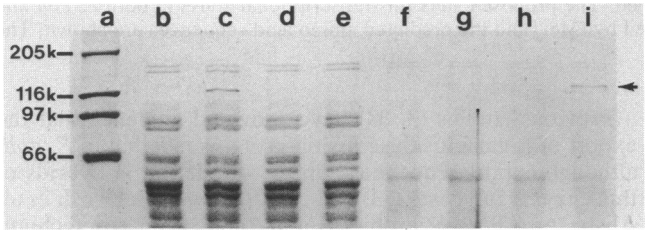

FIG. 2. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of cytoplasmic and membrane fractions from *E. coli*. Cytoplasmic and membrane fractions were prepared by the freeze-thaw protocol described for the localization of β-galactosidase activity. Cytoplasmic lanes (b to e) contained approximately 7 μg of protein per lane; membrane lanes (f to i) contained approximately 3 μg of protein per lane. A mixture of proteins (SDS-6H; Sigma Chemical Co., St. Louis, Mo.) was run in lane a (myosin, β-galactosidase, phosphorylase *b*, and bovine serum albumin). Other lanes contained fractions from strain TC4 harboring the following plasmids: pLOI193 (lanes b and f), pLOI301 (lanes c and g), pLOI303 (lanes d and h), pLOI302 (lanes e and i). The arrow indicates the region of the hybrid β-galactosidase proteins.

Figure 4 shows a comparative plot of hydropathy for the amino-terminal ends of the three β-galactosidase hybrids. From this plot, the hydrophobic nature of pLOI302 and hydrophilic nature of pLOI301 are readily apparent. The hybrid encoded by pLOI303 is intermediate, with a moderately hydrophobic N terminus, and exhibits partial membrane localization.

Transcriptional initiation sites in *E. coli* and *Z. mobilis*. The position of transcriptional initiation in three hybrid β-galactosidase genes (pLOI301, pLOI302, and pLOI303) was examined by primer extension analysis and S1 nuclease mapping in both *E. coli* and *Z. mobilis*. S1 nuclease mapping of RNAs from strains carrying pLOI301 indicated the presence of the same major transcriptional initiation site in both *E. coli* and *Z. mobilis* (Fig. 5). This major band was located 18 base pairs upstream from the in-frame methionine and has been labeled bp 1 in Fig. 3A and 5. Fainter bands below may represent degradation products. Primer extension analyses of this plasmid were very ambiguous and contained many additional bands in preparations from both organisms (not shown).

The results of primer extension analyses and S1 mapping were in general agreement for both pLOI302 and pLOI303. Figure 6 shows the results of primer extension experiments with *E. coli* and *Z. mobilis* containing pLOI302. In *Z. mobilis*, transcription appeared to be initiated from two positions separated by a single base, suggestive of microheterogeneity in initiation. The first of these has been labeled bp 1 in Fig. 3B and 6. Fainter bands below may represent degradation products. The results for *E. coli* carrying pLOI302 were similar, except that the lower bands observed in *Z. mobilis* were more pronounced in *E. coli*. Two major sites appeared to be used for transcriptional initiation, at positions −1 and 11. These two major bands were also observed in S1 mapping experiments, with lighter bands below (not shown). The fainter bands probably represent degradation products.

Primer extension analyses of RNAs prepared from *E. coli* and *Z. mobilis* containing pLOI303 are shown in Fig. 7. In *Z. mobilis*, transcription was initiated from a cytosine residue (band labeled bp 1 in Fig. 3C and 7), followed by a band of lower intensity at position 4. *E. coli* exhibited bands of high intensity at positions −9 (adenine residue), −3 (thymine residue), and +4 (cytosine residue). These multiple bands were also evident in S1 mapping experiments and may represent multiple sites of transcriptional initiation in *E. coli*.

DISCUSSION

We have constructed a number of *lacZ'* gene fusions with N termini from *Z. mobilis*, which simultaneously provided

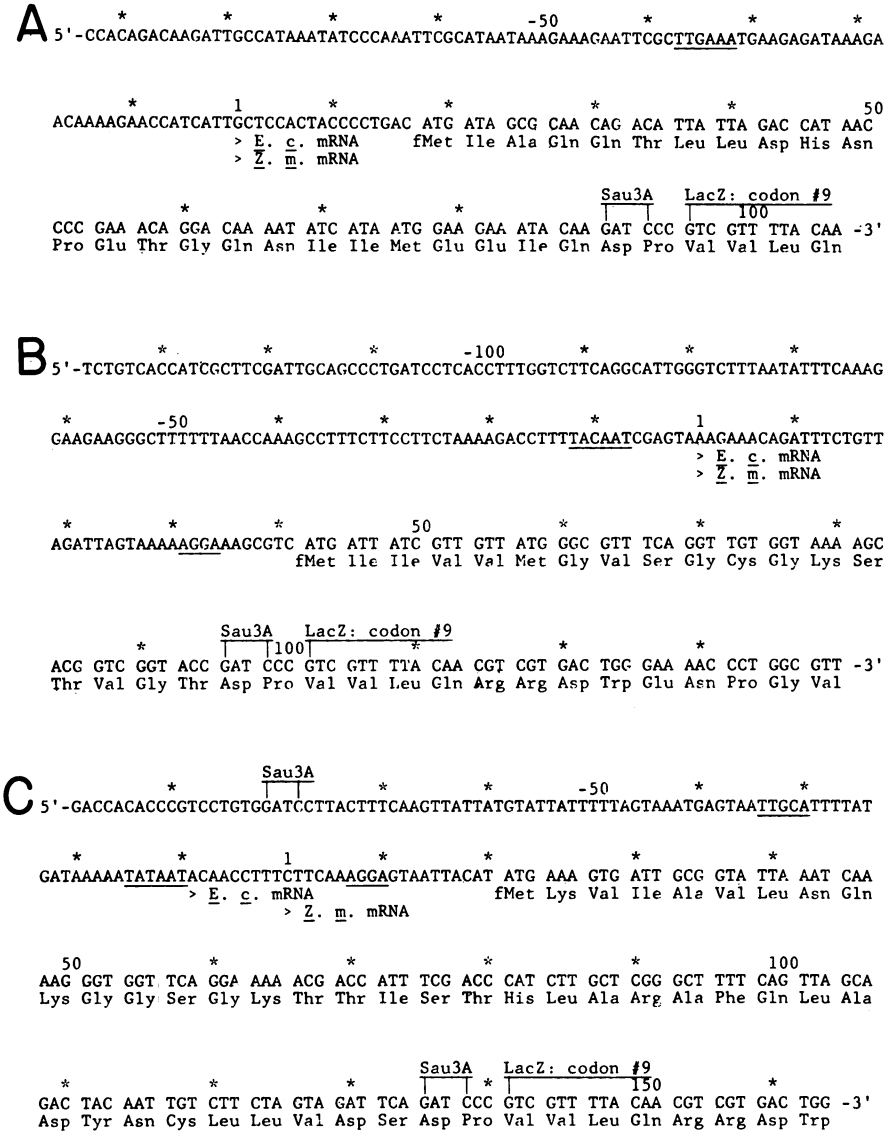

FIG. 3. Nucleotide sequences of the *Z. mobilis* DNA inserts from three β-galactosidase fusions. (A) pLOI301; (B) pLOI302; (C) pLOI303. The presumed −35, −10, and ribosomal-binding sequences are underlined. The proposed sites for transcriptional starts in both *E. coli* and *Z. mobilis* are marked (>). The proposed translational starts are indicated by fMet, and the predicted amino acid sequences are shown. The *Sau*3A sites indicate the point of fusion with the truncated *lacZ'* gene.

β-galactosidase activity and circumvented the need for a lactose permease gene in *E. coli*. It was originally anticipated that these functional gene fusions would be of two principal types: exported fusion genes (outer surface of the plasma membrane, periplasmic region, and outer membrane) and fusion genes with very high levels of β-galactosidase activity (34). None of the gene fusions examined exported a functional hybrid protein, but all exhibited high levels of activity. Over half of these fusion strains produced functional hybrid β-galactosidase, which remained associated with the membrane fraction after cell lysis. The associations of these hybrid proteins with the membrane may increase cell permeability and contribute to the observed lactose-positive phenotype in the absence of a functional permease. Weaker membrane associations may also occur with the other hybrid enzymes, although these enzymes did not sediment with the membrane fraction under our conditions.

Previous studies (3, 34) have provided evidence that the export of β-galactosidase fusion proteins is lethal in *E. coli*, although export sequences appear to function in hybrids of this enzyme in yeasts (33) and in *S. lividans* (Eckhardt et al. Abstr. Am. Soc. Microbiol. First Annu. Conf. Biotechnol. 1986). Such lethality is consistent with the lack of exported fusion genes from our selection procedure in *E. coli*.

The gene fusions which were membrane associated in *E. coli* were also membrane associated in *Z. mobilis*, indicating a similarity of requirements for anchor sequences and translational initiation sites, despite pronounced differences in membrane composition (6, 15). Comparison of the sequences from a membrane-associated fusion product (pLOI302), from a partially membrane-localized fusion (pLOI303), and from a soluble fusion product (pLOI301) indicate that five to eight consecutive, hydrophobic amino acids near the N terminus can serve to anchor a functional

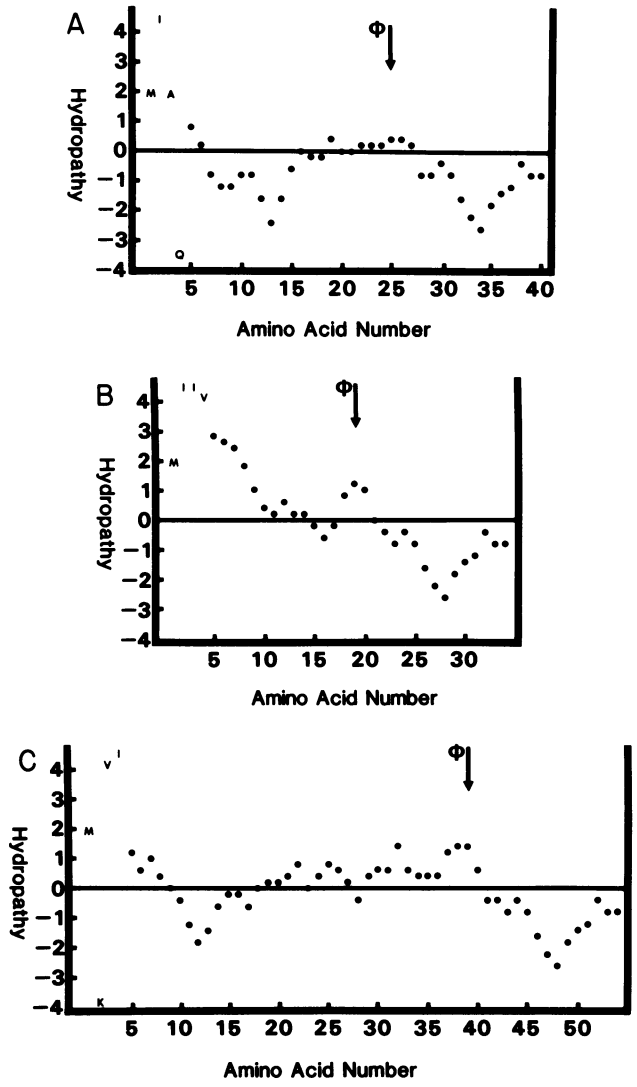

FIG. 4. Hydropathy plots of β-galactosidase fusion proteins. Hydropathy was determined by using an averaging window of 9 amino acids. (A) pLOI301; (B) pLOI302; (C) pLOI303. Arrows indicate points of fusion of *Z. mobilis* DNA in frame with *lacZ'*. Letters plotted at first four positions correspond to the amino acids at these positions in the protein (single-letter amino acid code) which otherwise would not be plotted with an averaging window of 9. They are positioned according to their respective hydropathy values.

β-galactosidase of over 120,000 daltons to the cell membrane in *Z. mobilis* and *E. coli*. This result compares favorably with other studies of anchor sequence requirements in mammalian and microbial systems (1, 10, 24, 28).

A comparison of the transcriptional initiation sites of fusion genes in *Z. mobilis* and *E. coli* indicated that both organisms generally recognized similar regions of DNA. However, the recognition of sequences was not always identical. *Z. mobilis* exhibited one or two closely spaced primary transcriptional starts, while *E. coli* initiated high levels of transcription at additional sites in two of the three promoters studied. It is likely that this increased transcription and higher number of plasmid copies in *E. coli* are in large part responsible for the higher levels of hybrid enzyme activity observed in *E. coli*.

In general, the *Z. mobilis* fragments examined share many of the features described for the promoter regions of other procaryotes (14, 16, 30). However, these fragments were originally selected as providing functional promoters in *E. coli* and may not be representative for *Z. mobilis*. The *Z. mobilis* DNA fragment in pLOI301 did not contain a sequence characteristic of that found in the −10 region of *E. coli* (14, 30) but was very rich in adenine. The pLOI301 fragment contained a near-consensus sequence, TTGAAA, in the −35 region (14, 30). The *Z. mobilis* fragment in pLOI302 contained the sequence TACAAT in the −10 region, closely resembling the consensus sequence proposed for *E. coli* in this region (14, 30). It did not contain a recognizable sequence characteristic of the −35 region. The sequence of the pLOI303 fragment most closely resembled the consensus promoter proposed for *E. coli* (14, 30) and contained the sequence TATAAT in the −10 region and TTGCA in the −35 region, with other possible promoters further upstream. It is interesting that the levels of expression of this latter fusion gene (pLOI303) were particularly high in both *E. coli* and *Z. mobilis*.

A potential ribosomal-binding site (16), AGGA, was identified 8 to 12 base pairs upstream from an in-frame start codon in all three fusions. This sequence was 20 to 75 base pairs below the primary initiation of transcription identified in *Z. mobilis*. Palindromic sequences rich in A+T were identified immediately upstream from the major transcriptional initiation sites in *Z. mobilis* for two of the fusion genes. In pLOI302, a 6-base-pair palindrome (TAAAAG-AC-CTTTTA) is separated by two base pairs and runs into the −10 region of the promoter. In pLOI303, an analogous 5-base-pair palindrome (ATTTT-ATATA-AAAAT) is separated by five base pairs and also runs into the −10 region of

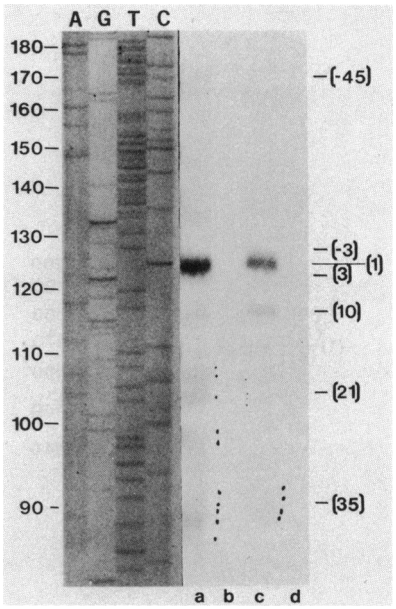

FIG. 5. S1 nuclease mapping of the start of transcription of the β-galactosidase fusion gene of pLOI301. A sequencing ladder was generated by collapsed-plasmid sequencing (A, G, T, and C) and is shown with numbers on the left indicating the sizes of DNA fragment in base pairs, including the primer. Lanes: a, *E. coli* containing pLOI301; b, *E. coli* control; c, *Z. mobilis* containing pLOI301; d, *Z. mobilis* control. Numbers in parentheses on the right correspond to the numbering of nucleotides in Fig. 3A. The most likely transcriptional start in both organisms has been designated by the number 1. Black dots adjacent to S1 nuclease bands are due to the overzealous use of a permanent marker by the investigators.

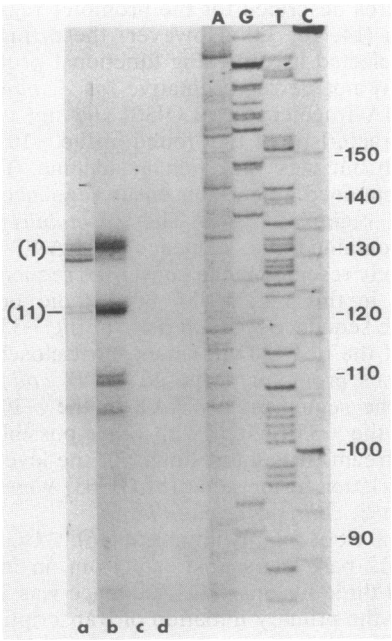

FIG. 6. Primer extension analysis of transcriptional start of the β-galactosidase fusion gene on pLOI302. The sequence ladder was generated by collapsed-plasmid sequencing (A, G, T, and C). The sizes of the DNA fragments, including the primer, are given in base pairs on the right. Lanes: a, *Z. mobilis* containing pLOI302; b, *E. coli* containing pLOI302; c, *Z. mobilis* control; d, *E. coli* control. Numbers in parentheses on the left correspond to numbering in Fig. 3B. The most likely transcriptional start in both organisms has been designated by the number 1.

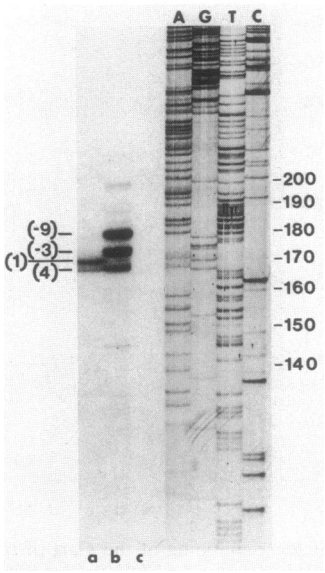

FIG. 7. Primer extension analysis of the start of transcription of the β-galactosidase fusion gene on pLOI303. The sequence ladder was generated by collapsed-plasmid sequencing (A, G, T, and C). The sizes of the DNA fragments, including the primer, are given in base pairs on the right. Lanes: a, *Z. mobilis* containing pLOI303; b, *E. coli* containing pLOI303; c, *E. coli* control. The *Z. mobilis* control contained no bands (not shown). Numbers in parentheses on the left correspond to nucleotide numbering in Fig. 3C. The most likely start of transcription in *Z. mobilis* has been designated by the number 1.

the promoter. Repeated sequences were identified in pLOI301 in the A+T-rich region immediately above the major site of transcription initiation in *Z. mobilis*, although it was not possible to identify a sequence resembling that proposed for the −10 region of *E. coli*. However, the sequence ATAAGA occurred twice in this general region of pLOI301, along with the nearly matched sequences AAAAGA and ATAAATA. It is possible that both the repeated sequences in pLOI301 and the palindromes in pLOI302 and pLOI303 serve as important recognition sites for transcription or as terminators of upstream genes. Although the promoter recognition sequences and ribosomal-binding sequences may not be identical in *Z. mobilis* and *E. coli*, the results of this study clearly indicate that similar sequences can function as promoters in both organisms.

ACKNOWLEDGMENTS

We thank K. T. Shanmugam and W. E. Gurley for their review of this manuscript.

This research was supported in part by the Florida Agricultural Experiment Station and by grants from the National Science Foundation (DMB 8204928), the Department of Energy, Office of Basic Energy Science (FG05-86ER3574), and the Department of Agriculture, Alcohol Fuels Program (86-CRCR-1-2134). We gratefully acknowledge support by the Egyptian Ministry Higher of Education to Y. A. Osman.

LITERATURE CITED

1. **Adams, G. A., and J. K. Rose.** 1985. Structural requirements of a membrane-spanning domain for protein anchoring and cell surface transport. Cell **41**:1007–1015.
2. **Bagdasarian, M., and K. N. Timmis.** 1982. Host-vector systems for gene cloning in *Pseudomonas*. Curr. Top. Microbiol. Immunol. **96**:47–67.
3. **Benson, S. A., M. N. Hall, and T. J. Silhavy.** 1985. Genetic analysis of protein export in *Escherichia coli* K12. Annu. Rev. Biochem. **54**:101–134.
4. **Bialkowska-Hobrzanska, H., C. A. Gilchrist, and D. T. Denhardt.** 1985. *Escherichia coli rep* gene: identification of the promoter and N terminus of the Rep protein. J. Bacteriol. **164**:1004–1010.
5. **Byung, M. O.-K., J. B. Kaper, and L. O. Ingram.** 1986. Construction of a new vector for the expression of foreign genes in *Zymomonas mobilis*. J. Ind. Microbiol.**1**:9–15.
6. **Carey, V. C., and L. O. Ingram.** 1983. Lipid composition of *Zymomonas mobilis*: effects of ethanol and glucose. J. Bacteriol. **154**:1291–1399.
7. **Carey, V. C., S. K. Walia, and L. O. Ingram.** 1983. Expression of a lactose transposon (Tn*951*) in *Zymomonas mobilis*. Appl. Environ. Microbiol. **46**:1163–1168.
8. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:165–170.
9. **Conway, T., M. O.-K. Byun, and L. O. Ingram.** 1987. Expression vector for *Zymomonas mobilis*. Appl. Environ. Microbiol. **53**:235–241.
10. **Davis, N. G., and P. Model.** 1985. An artificial anchor domain: hydrophobicity suffices to stop transfer. Cell **41**:607–614.
11. **Favaloro, J., R. Treisman, and R. Kamen.** 1980. Transcription maps of polyoma virus-specific RNA: analysis by two-dimensional nuclease S1 gel mapping. Methods Enzymol. **65**:718–749.
12. **Fried, V. A., and A. Novick.** 1973. Organic solvents as probes for the structure and function of the bacterial membrane: effects of ethanol on the wild type and an ethanol-resistant mutant of *Escherichia coli* K-12. J. Bacteriol. **114**:239–248.
13. **Guerry, P., J. van Embden, and S. Falkow.** 1974. Molecular nature of two nonconjugative plasmids carrying drug resistance genes. J. Bacteriol. **117**:619–630.
14. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and

analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
15. **Ingram, L. O., and T. M. Buttke.** 1984. Effect of alcohols on microorganisms. Adv. Microb. Physiol. **25**:256–300.
16. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**:1–45.
17. **Kyte, J., and R. F. Doolittle.** 1982. A simple method of displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
18. **Luria, S. E., and M. Delbruck.** 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28**:491–511.
19. **Mandel, M., and A. Higa.** 1970. Calcium dependent bacteriophage DNA infection. J. Mol. Biol. **53**:159–162.
20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Michaelis, S., and J. Beckwith.** 1982. Mechanism of incorporation of cell envelope proteins in *Escherichia coli*. Annu. Rev. Microbiol. **36**:435–465.
22. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Montenecourte, B. S.** 1985. *Zymomonas*, a unique genus of bacteria, p. 261–289. *In* A. L. Demain and N. A. Solomon (ed.), Biology of industrial microorganisms. The Benjamin-Cummings Publishing Co., Inc., Menlo Park, Calif.
24. **Oliver, D.** 1985. Protein secretion in *Escherichia coli*. Annu. Rev. Microbiol. **39**:615–648.
25. **Proudfoot, N. J., M. H. M. Shander, J. L. Manley, M. L. Gefter, and T. Maniatis.** 1980. Structure and *in vitro* transcription of human globin genes. Science **209**:1329–1336.
26. **Pustell, J., and F. C. Kafatos.** 1984. A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis, and homology determination. Nucleic Acids Res. **12**:643–655.
27. **Randall, L. L.** 1983. Translocation of domains of nascent periplasmic proteins across the cytoplasmic membrane is independent of elongation. Cell **33**:231–240.
28. **Randall, L. L., and S. J. S. Hardy.** 1984. Export of protein in bacteria. Microbiol. Rev. **48**:290–298.
29. **Rogers, P. L., K. L. Lee, M. L. Skotnicki, and D. E. Tribe.** 1982. Ethanol production by *Zymomonas mobilis*. Adv. Biochem. Eng. **23**:37–84.
30. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–353.
31. **Rossi, J. J., X. Soberon, Y. Marumoto, J. McMahon, and K. Itakura.** 1983. Biological expression of an *Escherichia coli* consensus promoter and some mutant derivatives. Proc. Natl. Acad. Sci. USA **80**:3203–3207.
32. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
33. **Silhavy, T. J., and J. R. Beckwith.** 1985. Uses of *lac* fusions for the study of biological problems. Microbiol. Rev. **49**:398–418.
34. **Silhavy, T. J., S. A. Benson, and S. D. Emr.** 1983. Mechanism of protein localization. Microbiol. Rev. **47**:313–344.
35. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36. **Simon, R., U. Priefer, and A. Puhler.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram-negative bacteria. Biotechnology **1**:784–791.
37. **Skotnicki, M. L., D. E. Tribe, and P. L. Rogers.** 1980. R-plasmid transfer in *Zymomonas mobilis*. Appl. Environ. Microbiol. **41**:889–893.
38. **Stokes, H. W., E. L. Dally, M. D. Yablonsky, and D. E. Eveleigh.** 1983. Comparison of plasmids in strains of *Zymomonas mobilis*. Plasmid **9**:138–146.
39. **Swings, J., and J. DeLey.** 1977. The biology of *Zymomonas*. Bacteriol. Rev. **41**:1–46.
40. **Weinstock, G., C. ap Rhys, M. Berman, B. Hampar, D. Jackson, T. Silhavy, J. Weisman, 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.
41. **Williams, S. A., B. E. Slatko, L. S. Moran, and S. M. DeSimone.** 1986. Sequencing in the fast lane: a rapid protocol for [α-^{35}S]ATP dideoxy DNA sequencing. Biotechnology **4**:138–147.
42. **Willis, R. C., R. G. Cirakoglu, C. Morris, G. D. Schellenberg, N. H. Gerber, and C. E. Furlong.** 1974. Preparation of the periplasmic binding proteins from *Salmonella typhimurium* and *Escherichia coli*. Arch. Biochem. Biophys. **161**:64–75.