Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 1101–1105, February 1983
Microbiology

Chromogenic identification of genetic regulatory signals in *Bacillus subtilis* based on expression of a cloned *Pseudomonas* gene

(heterospecific gene expression/catechol 2,3-dioxygenase/TOL *xylE* gene/expression vector)

MARK M. ZUKOWSKI, DAIRENA F. GAFFNEY*, DENIS SPECK, MURIEL KAUFFMANN, ANNIE FINDELI, ANNE WISECUP, AND JEAN-PIERRE LECOCQ

Transgène S.A, 11, rue Humann, 67000 Strasbourg, France

Communicated by Sydney Brenner, November 22, 1982

ABSTRACT **A method to isolate fragments of DNA that promote gene expression in *Bacillus subtilis* is described. The system is based on production of catechol 2,3-dioxygenase [$CatO_2$ase; catechol:oxygen 2,3-oxidoreductase (decyclizing), EC 1.13.11.2] encoded by the *Pseudomonas putida* TOL plasmid gene *xylE*. The gene was transferred to a *B. subtilis*/*Escherichia coli* plasmid vector to construct pTG402. Although *xylE* is functionally expressed in *E. coli*, $CatO_2$ase is not detected in *B. subtilis* unless a fragment of DNA capable of promoting gene expression is ligated into a cleavage site on pTG402 upstream from *xylE*. Fragments of chromosomal DNA from *B. subtilis*, *Bacillus licheniformis*, *Bacillus pumilus*, and *E. coli* are shown to promote *xylE* gene expression in *B. subtilis*. The special feature of the system is the method of detection: colonies of cells that express *xylE* become yellow within seconds after selection plates are sprayed with catechol, a colorless substrate that is converted by $CatO_2$ase to the yellow product, 2-hydroxymuconic semialdehyde. The complete nucleotide sequence of *xylE* is presented. Strong complementarity between the ribosome binding site and 16S rRNA suggests that *xylE* mRNA translation in *B. subtilis* may commence at the same site as that recognized by *P. putida*. Identity of $CatO_2$ase produced in *B. subtilis*, *E. coli*, and *P. putida* support the hypothesis. Our sensitive color assay offers an approach to develop plasmid gene expression vectors for a wide variety of host organisms.**

Bacillus subtilis is an attractive alternative to *Escherichia coli* as a host for expression of cloned genes. The Gram-positive organism is nonpathogenic, free of endotoxins, and an important producer of extracellular enzymes on a large industrial scale. Critical to the development of the microorganism as a host–vector system for recombinant DNA technology is the efficient expression of heterospecific genes. To express plasmid-borne genes in *B. subtilis*, transcriptional or translational signals that differ from those of *E. coli* are required (1). Plasmid vectors suitable for cloning fragments of DNA that carry transcriptional promoter or termination signals for Gram-negative bacteria into *E. coli* have been characterized (2–6). Detection in these systems is based on expression of genes that encode β-galactosidase (4–6) or confer antibiotic resistance to host cells (2, 3). An approach similar to the latter has been successful in *B. subtilis* using chloramphenicol acetyltransferase genes originating from *Bacillus pumilus* (7) or the transposable genetic element Tn9 (8). In the *E. coli* β-galactosidase system, selection of DNA fragments that promote expression of the *lacZ* gene is based on an easily visualized color change of bacterial colonies grown on indicator plates containing a chromogenic substrate (4). An analogous system that functions in *B. subtilis* would greatly facilitate the effort to decipher problems of heterospecific gene expression in Gram-positive bacteria.

In this report, we present a method whereby fragments of DNA that promote expression of a foreign gene in *B. subtilis* are detected by a change of color of bacterial colonies. The system is based on the cloning and expression, in *B. subtilis*, of the *xylE* gene, which originated from the TOL plasmid pWWO (9) of *Pseudomonas putida* mt-2. The assay is rapid and inexpensive, does not require special indicator plates but offers the advantages of a genetic indicator test (10), and can be used for the development of efficient plasmid gene expression vectors.

MATERIAL AND METHODS

Bacterial Strains and Plasmids. The *B. subtilis* strains used are derivatives of Marburg strain 168. Strains BZ2 *cysB3 recE4* and TGB1 *trpC2 recE4 spo331* were constructed by transformation (11). MI112 *arg15 leuB thr5* $r_M^- m_M^-$ *recE4* was from T. Tanaka. *Bacillus licheniformis* 9945A and *Bacillus pumilus* BP1 were obtained from the Bacillus Genetic Stock Center (Ohio State University, Columbus). *E. coli* strain BZ18 was from W. Arber; C600 $r_k^- m_k^+$ was from J. W. Little; *Pseudomonas putida* mt-2 was donated by K. Timmis. The bifunctional *E. coli*/*B. subtilis* plasmid pHV33 (12) was obtained from R. Dedonder. Plasmid DNA was prepared by an alkaline extraction procedure (13) or a cleared lysate method (14) followed by cesium chloride/ethidium bromide density gradient centrifugation (15).

Transformations. Plasmid DNA was introduced into cryogenically preserved *E. coli* (16) by the transformation method of Lederberg and Cohen (17). Transformants were selected by antibiotic resistance (ampicillin, Amp, 50 μg/ml; tetracycline, Tet, 15 μg/ml; chloramphenicol, Cml, 25 μg/ml) of host cells. Plasmids were introduced into *B. subtilis* by transformation of competent cells (11, 18) or protoplasts (19) and their presence was selected by resistance of host cells to Cml (10 μg/ml).

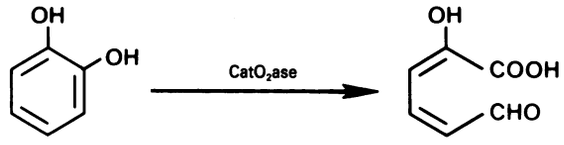
Construction of Recombinant Plasmids. Restriction endonucleases were purchased from New England BioLabs and Bethesda Research Laboratories. Digestions were carried out according to the recommendations of the supplier. To construct pTG402-derivative plasmids that express *xylE* in *B. subtilis*, purified chromosomal DNA (20) was digested to completion with *Sau*3A and then mixed at a 2:1 mass ratio, chromosomal DNA to plasmid DNA, with pTG402 previously cleaved at the unique *Bam*HI site. DNA fragments were ligated overnight at 15°C in 66 mM Tris·HCl, pH 7.9/6.6 mM $MgCl_2$/10 mM dithiothreitol/0.5 mM ATP with T4 DNA ligase (Boehringer Mannheim).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: Amp, ampicillin; Tet, tetracycline; Cml, chloramphenicol; $CatO_2$ase, catechol 2,3-dioxygenase.

* Present address: Institute of Virology, Church St., Glasgow G11 5JR, U.K.

Assay for Catechol 2,3-Dioxygenase Activity. Functional expression of the *xylE* gene was detected by spraying antibiotic selection plates with an aqueous solution of 0.5 M catechol. Colonies of cells that express the gene become yellow due to the conversion by catechol 2,3-dioxygenase [$CatO_2$ase; catechol:oxygen 2,3-oxidoreductase (decyclizing), EC 1.13.11.2] of catechol to 2-hydroxymuconic semialdehyde (21–23) as shown below.



For enzymatic assays, cell cultures were grown overnight at 37°C in 10 ml of Penassay broth (Difco). Cells were washed with 20 mM phosphate buffer (pH 7.2), suspended in 100 mM phosphate buffer, pH 7.5/10% (vol/vol) acetone (AP buffer), and broken by sonication for 3 min. Extracts were centrifuged in an Eppendorf Microfuge for 15 min at 4°C to remove cellular debris. Supernatants were diluted with AP buffer or assayed directly. $CatO_2$ase specific activities were determined spectrophotometrically as described (24). One milliunit corresponds to the formation at 30°C of 1 nmol of 2-hydroxymuconic semialdehyde per min. Protein concentrations were measured by the method of Lowry *et al.* (25), using bovine serum albumin (Sigma) as reference.

Sequence Analysis. Fragments of DNA for sequence analysis were subcloned onto derivatives of bacteriophage M13 mp7 (26) and M13 mp701 (D. R. Bentley, personal communication). The dideoxy chain-termination method of sequence analysis (27) was used throughout.

RESULTS

Construction of the Plasmid Vector pTG402. The *xylE* gene has previously been shown to reside on the *Xho* I fragment of pWWO (28). Further analysis showed that the gene was wholly contained within a 2.0-kilobase subfragment of the *Xho* I fragment that carries a *Bam*HI site at one extremity and a *Xho* I site at the other (21). Transcription of *xylE* mRNA has been shown to begin proximal to the *Bam*HI site (22) but the gene, when cloned in *E. coli* (22, 23), appears to require a vector promoter for efficient expression. The *Bam*HI/*Xho* I fragment carrying *xylE* was first transferred from pWWO to pBR322 (29) by ligation of the fragment into the large segment of pBR322 that remained after double digestion with *Bam*HI and *Sal* I (30). The resulting plasmid, pTG206 (Fig. 1), was introduced into *E. coli* BZ18 by transformation. Expression of *xylE*, presumably under control of the active *tet* promoter on pBR322, was detected by spraying Amp-resistant (Amp^r) Tet-sensitive (Tet^s) colonies with catechol. Yellow colonies indicated the presence of $CatO_2$ase; cells that did not express *xylE* remained white.

The gene was next transferred from pTG206 to pHV33 (Fig. 1), a plasmid vector that functions in *E. coli* or *B. subtilis* (12). The newly constructed plasmid was selected in *E. coli* BZ18 by spraying Amp^r Cml-resistant (Cml^r) Tet^s transformant colonies with catechol. Plasmid DNA originating from one of the yellow colonies was purified and then digested with restriction endonucleases to construct a physical map of pTG402 (Fig. 1).

pTG402 plasmid was introduced into *B. subtilis* TGB1 by transformation of competent cells. Transformants selected for Cml^r after 24 hr of incubation at 37°C were sprayed with catechol. Yellow coloration failed to develop.

To determine whether rearrangement or deletion of the *xylE* gene was responsible for lack of expression in *B. subtilis*,

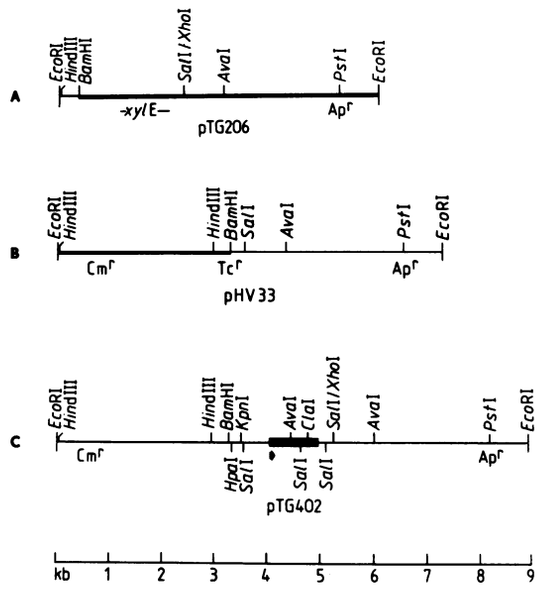


FIG. 1. (*A*) Partial restriction endonuclease map of plasmid pTG206. The heavy line represents the portion of pTG206 that contributed to the construction of pTG402. (*B*) Physical map of pHV33, as described elsewhere (12). *E. coli* host cells bearing pHV33 are Amp^r (50 μg/ml), Tet^r (15 μg/ml), and Cml^r (25 μg/ml), whereas only Cml^r (10 μg/ml) is expressed in *B. subtilis*. The heavy line represents the portion of pHV33 that contributed to the construction of pTG402. (*C*) Partial restriction endonuclease map of pTG402. Plasmids pTG206 and pHV33 were digested with *Eco*RI/*Bam*HI. Fragments were ligated with T4 DNA ligase and pTG402 was selected as described in the text. The very heavy line represents *xylE*, and the arrow indicates direction of mRNA transcription (22). The *xylE* gene on pTG402 is expressed in *E. coli* but not in *B. subtilis* until a sequence of DNA that promotes expression is inserted into a cleavage site, such as *Bam*HI, upstream from the structural gene. kb, Kilobase(s).

pTG402 was purified from TGB1 and then reintroduced into *E. coli* BZ18. All Amp^r Cml^r Tet^s transformant colonies became yellow when sprayed with catechol. This showed that the *xylE* gene had remained intact in the *Bacillus* host in which gene expression was not detected. The absence of $CatO_2$ase suggested a lack of promoter activity for the *xylE* gene in the Gram-positive host.

Expression of the *xylE* Gene in *B. subtilis*. Sequences of DNA that promote expression of the *xylE* gene on pTG402 were isolated as follows. Chromosomal DNA from *B. subtilis* BZ2 was digested to completion with *Sau*3A. The fragments were ligated *in vitro* to *Bam*HI-digested pTG402. Ligation products were introduced into *E. coli* BZ18. Propagation of plasmids in *E. coli* ensured the formation of multimeric molecules necessary for subsequent transformation into *B. subtilis* (31). *E. coli* transformants were selected by overnight incubation at 37°C in 20 ml of L broth supplemented Amp at 50 μg/ml. Plasmid DNA was isolated from the mixed cell population by a cleared lysate method (14). A small portion of the DNA from the cleared lysate was then used as donor DNA for transformation of *B. subtilis* TGB1 or MI112 recipient cells. It has been suggested that transformation efficiencies can be increased by using the restriction/modification-defective strain MI112 (32), but in our hands transformation frequencies in TGB1 and MI112 were similar (data not shown). Cml^r plasmid-bearing *B. subtilis* transformants were sprayed with catechol. The procedure resulted in the isolation of yellow colonies at a frequency of 4.4×10^{-3} of total transformants. This method, however, could yield plasmids identical in origin due to the nature of the donor DNA used in transformation of *B. subtilis*. We therefore pursued an alter-

native method that greatly reduced the possibility of selecting identical recombinant plasmids. Ligation products were introduced directly into *B. subtilis* by protoplast transformation (19) in which multimeric forms of plasmid DNA are not required for efficient uptake and maintenance. Selection was for Cml^r colonies that arose after 48 hr of incubation at 37°C on regeneration medium supplemented with Cml at 10 μg/ml. When transformants were sprayed with catechol, yellow colonies were observed at a frequency of 6.8×10^{-3} of total transformants.

To examine whether fragments of DNA from other species might also promote expression of *xylE*, chromosomal DNA was prepared from *B. licheniformis* 9945A, *B. pumilus* BP1, and *E. coli* C600 $r_k^- m_k^+$. DNA was digested to completion with *Sau*3A, ligated into the *Bam*HI site of pTG402, and used to transform *B. subtilis* MI112 protoplasts. The frequency of Cml^r yellow colonies obtained from each *Bacillus* donor approximated that observed with the *B. subtilis* chromosomal DNA. The *E. coli* donor DNA, however, gave rise to only 2 yellow colonies from more than 10^4 total Cml^r transformants.

$CatO_2$ase Activities. Plasmids from several transformants that expressed *xylE* were examined for DNA insert size and specific activity of the $CatO_2$ase produced (Table 1). Enzyme activities were measured on duplicate cell-free extracts. This provided an estimate of the level of gene expression as measured by the quantity or efficiency of the translation product. Neither pHV33 nor pTG402 had any $CatO_2$ase activity in *B. subtilis*, but pTG402 derivatives that expressed *xylE* displayed a wide range of activities. The level of $CatO_2$ase produced in *B. subtilis* compared favorably, in most instances, with that of the uninduced enzyme in *P. putida* mt-2. There is no correlation between the size of the insert and the activity expressed. Plasmid copy numbers, although not determined for each case, are most likely similar because the plasmids display relatedness in structure and are consistent in quantity when purified.

Table 1. $CatO_2$ase specific activity in *B. subtilis*

Plasmid	Insert size, base pairs	Specific activity, milliunits/mg
None		0
pHV33		0
pTG402		0
pWWO*		163
	B. subtilis insert	
pTG403	500	238
pTG404	140	21
pTG405	575	1,703
pTG406	950	124
pTG407	335	374
pTG408	500	1,175
pTG409	900	170
pTG410	500	962
pTG411	100	58
pTG413	500	14
pTG414	1,500	288
pTG416	1,100	2,011
pTG417	900	1,853
pTG418	700	217
pTG420	600	551
pTG422	600	192
	B. pumilus insert	
pTG435	200	1,794
pTG436	550	327
pTG437	1,800	725
	B. licheniformis insert	
pTG438	700	555
pTG439	400	2,240
pTG440	600	1,018
pTG441	500	2,727
	E. coli insert	
pTG432	600	42
pTG433	500	324

*Sau*3A fragments of chromosomal DNA were ligated into the *Bam*HI site of pTG402. Insert size was determined by agarose gel electrophoresis of fragments produced by digestion with *Cla* I, *Hin*dIII/*Ava* I, or *Bam*HI/*Cla* I. Bacteriophage λcI857*Sam*7 DNA digested with *Hin*dIII or *Eco*RI/*Hin*dIII was used as the standard for molecular weight determinations.

* Host cells, *P. putida* mt-2, were grown overnight in 10 ml of L broth at 30°C. Enzyme activities were calculated exactly as described for *B. subtilis*.

Identity of $CatO_2$ase in *P. putida*, *E. coli*, and *B. subtilis*. To show that the enzyme produced in *B. subtilis* is the same as that produced in *P. putida* mt-2 or *E. coli* BZ18/pTG206, partially purified cell-free extracts from each organism were resolved by polyacrylamide gel electrophoresis (Fig. 2). The nondenaturing conditions used in the gel system allowed rapid identification of $CatO_2$ase by spraying the gel with catechol after electrophoresis was terminated. Yellow bands appeared at coincident locations in the gel for all extracts except that of the *B. subtilis* strain (TGB1/pTG402) that carried the *xylE* gene without a functional promoter. The identity of $CatO_2$ase electrophoretic mobilities suggests that enzyme produced in *B. subtilis* is not as fusion polypeptide but identical to the translation product observed in *Pseudomonas* and *E. coli*.

In an additional experiment, crude antiserum was prepared from a rabbit injected with $CatO_2$ase purified from *E. coli*. A double-diffusion Ouchterlony test showed that the antiserum

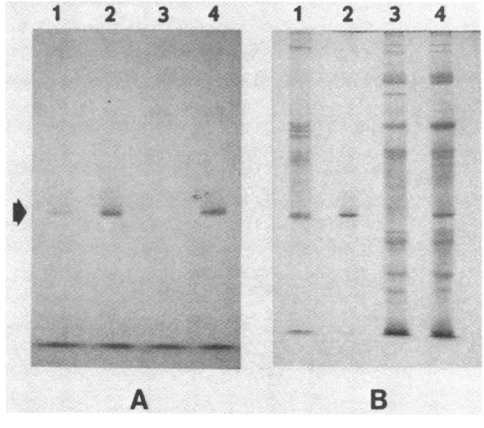


FIG. 2. Polyacrylamide gel electrophoresis (33) of $CatO_2$ase. Partially purified $CatO_2$ase preparations were obtained from stationary-phase cultures. Cell pellets were washed with 20 mM phosphate buffer (pH 7.2), suspended in AP buffer, and broken by sonication. Cellular debris was removed by centrifugation at 30,000 rpm for 1 hr at 4°C in a Beckman R65 Ti rotor and L8-70 ultracentrifuge. The supernatant was treated with acetone to a final concentration of 50% (vol/vol), proteins were allowed to precipitate at 4°C for 2 hr, and the mixture was centrifuged at 10,000 rpm for 10 min in a Sorvall HB-4 rotor and RC-5B centrifuge. Then, the supernatant was brought to 66% (vol/vol) acetone, proteins were allowed to precipitate for 2 hr at 4°C, and this mixture was centrifuged for 10 min at 10,000 rpm to give protein pellets that were resuspended in 1 ml of AP buffer. For electrophoresis, a 7.5% acrylamide separating gel and a 3.0% stacking gel were used. Protein (20–200 μg) was added to each lane so that total $CatO_2$ase activities (260 milliunits) were equal. For TGB1/pTG402, a protein quantity equivalent to that from TGB1/pTG403 was used. Electrophoresis was at a constant current of 25 mA for 5 hr. (*A*) The gel was sprayed with 0.5 M catechol. (*B*) The proteins were fixed, stained with Coomassie brilliant blue R250 (0.15%), and destained in 7.5% acetic acid/5% methanol (vol/vol). Extracts were from *P. putida* mt-2 (lane 1), *E. coli* BZ18/pTG206 (lane 2), *B. subtilis* TGB1/pTG402 (lane 3), and *B. subtilis* TGB1/pTG403 (lane 4). Arrow, $CatO_2$ase.

reacted with partially purified extracts from *P. putida* mt-2, *E. coli* BZ18/pTG206, and *B. subtilis* TGB1/pTG403 but did not cross-react with the *B. subtilis* extract prepared from TGB1/pTG402 (data not shown).

Nucleotide Sequences of the *xylE* Gene and the Ribosome Binding Site. The complete nucleotide sequence of *xylE* and the strategy used for its determination are detailed in Figs. 3 and 4, respectively. The location of the translational initiation codon and the reading frame presented are in perfect agreement with the sequence of the first nine amino acids determined from purified CatO_2ase. The gene encodes 306 amino acids that form a 35,047-dalton-subunit polypeptide. Since the active form of CatO_2ase has a molecular mass of 140,000 daltons (34), our results confirm previous reports that CatO_2ase consists of four identical subunits (35).

Extensive complementarity between the ribosome binding site (36) and the 3' region of 16S rRNA appears to be necessary for initiation of translation in *B. subtilis* (1, 37). From the proposed ribosome binding site for the *xylE* gene (Fig. 3), it can be predicted that ΔG, the free energy of base pairing (38) between this site and *B. subtilis* 16S rRNA, is -15.6 kcal (1 cal = 4.18 J). In addition, the distance between the ribosome binding site and the initiation codon is calculated to be nine bases. Both criteria fall well within the ranges of ΔG and spacer bases recently calculated from a series of ribosome binding sites for *B. subtilis* (37). The sequence data support the hypothesis that translational initiation of *xylE* mRNA in *B. subtilis* commences at the same initiation codon as that in *Pseudomonas* and *E. coli*.

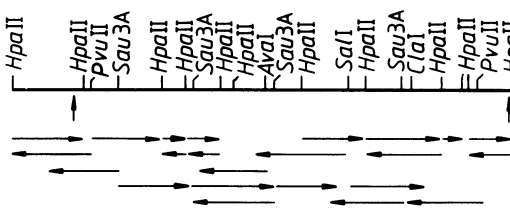


FIG. 4. Strategy used for subcloning and determining the sequences of the *xylE* gene fragments in bacteriophage M13 (horizontal arrows). The coding region of *xylE* is indicated by the vertical arrows.

DISCUSSION

We have described the construction and characterization of a system in which fragments of DNA that promote heterospecific gene expression in *B. subtilis* are detected by a simple color assay. Expression in *B. subtilis* of the *Pseudomonas* TOL plasmid gene *xylE*, which encodes the ring cleavage enzyme, CatO_2ase, serves as the indicator. When transferred to the plasmid cloning vector pTG402, *xylE* is expressed in *B. subtilis* only when a fragment of DNA capable of promoting gene expression is inserted into a cleavage site upstream from the gene. Because the enzyme catalyzes the conversion of a colorless substrate, catechol, to a yellow product, 2-hydroxymuconic semialdehyde, *xylE* gene expression is easily detected by spraying bacterial colonies that harbor plasmids of interest with an aqueous solution of catechol. Yellow coloration indicates that the gene is expressed. We have observed that the intensity of the color or

```
                             1                                                      50
GACAACATGAACTATGAAGAGGTGACGTC ATG AAC AAA GGT GTA ATG CGA CCG GGC CAT GTG CAG CTG CGT GTA CTG GAC ATG
                              Met Asn Lys Gly Val Met Arg Pro Gly His Val Gln Leu Arg Val Leu Asp Met

                                                         100
AGC AAG GCC CTG GAA CAC TAC GTC GAG TTG CTG GGC CTG ATC GAG ATG GAC CGT GAC GAC CAG GGC CGT GTC TAT CTG
Ser Lys Ala Leu Glu His Tyr Val Glu Leu Leu Gly Leu Ile Glu Met Asp Arg Asp Asp Gln Gly Arg Val Tyr Leu

                    150                                                       200
AAG GCT TGG ACC GAA GTG GAT AAG TTT TCC CTG GTG CTA CGC GAG GCT GAC GAG CCG GGC ATG GAT TTT ATG GGT TTC
Lys Ala Trp Thr Glu Val Asp Lys Phe Ser Leu Val Leu Arg Glu Ala Asp Glu Pro Gly Met Asp Phe Met Gly Phe

                                                 250
AAG GTT GTG GAT GAG GAT GCT CTC CGG CAA CTG GAG CGG GAT CTG ATG GCA TAT GGC TGT GCC GTT GAG CAG CTA CCC
Lys Val Val Asp Glu Asp Ala Leu Arg Gln Leu Glu Arg Asp Leu Met Ala Tyr Gly Cys Ala Val Glu Gln Leu Pro

            300                                                       350
GCA GGT GAA CTG AAC AGT TGT GGC CGG CGC GTG CGT TCC AGG CCC TCC GGG CAT CAC TTC GAG TTG TAT GCA GAC AAG
Ala Gly Glu Leu Asn Ser Cys Gly Arg Arg Val Arg Ser Arg Pro Ser Gly His His Phe Glu Leu Tyr Ala Asp Lys

                                         400
GAA TAT ACT GGA AAG TGG GGT TTG AAT GAC GTC AAT CCC GAG GCA TGG CCG CGC GAT CTG AAA GGT ATG GCG GCT GTG
Glu Tyr Thr Gly Lys Trp Gly Leu Asn Asp Val Asn Pro Glu Ala Trp Pro Arg Asp Leu Lys Gly Met Ala Ala Val

  450                                                           500
CGT TTC GAC CAC GCC CTC ATG TAT GGC GAC GAA TTG CCG GCG ACC TAT GAC CTG TTC ACC AAG GTG CTC GGT TTC TAT
Arg Phe Asp His Ala Leu Met Tyr Gly Asp Glu Leu Pro Ala Thr Tyr Asp Leu Phe Thr Lys Val Leu Gly Phe Tyr

                                550
CTG GCC GAA CAG GTG CTG GAC GAA AAT GGC ACG CGC GTC GCC CAG TTT CTC AGT CTG TCG ACC AAG GCC CAC GAC GTG
Leu Ala Glu Gln Val Leu Asp Glu Asn Gly Thr Arg Val Ala Gln Phe Leu Ser Leu Ser Thr Lys Ala His Asp Val

                                                          650
GCC TTC ATT CAC CAT CCG GAA AAA GGC CGC CTC CAT CAT GTG TCC TTC CAC CTC GAA ACC TGG GAA GAC TTG CTT CGC
Ala Phe Ile His His Pro Glu Lys Gly Arg Leu His His Val Ser Phe His Leu Glu Thr Trp Glu Asp Leu Leu Arg

                            700                                                             750
GCC GCC GAC CTG ATC TCC ATG ACC GAC ACA TCT ATC GAT ATC GGC CCA ACC CGC CAC GGC CTC ACT CAC GGC AAG ACC
Ala Ala Asp Leu Ile Ser Met Thr Asp Thr Ser Ile Asp Ile Gly Pro Thr Arg His Gly Leu Thr His Gly Lys Thr

                                                          800
ATC TAC TTC TTC GAC CCG TCC GGT AAC CGC AAC GAA GTG TTC TGC GGG GGA GAT TAC AAC TAC CCG GAC CAC AAA CCG
Ile Tyr Phe Phe Asp Pro Ser Gly Asn Arg Asn Glu Val Phe Cys Gly Gly Asp Tyr Asn Tyr Pro Asp His Lys Pro

                    850                                                           900
GTG ACC TGG ACC ACC GAC CAG CTG GGC AAA GCC TTC TTT TAC CAC GAC CGC ATT CTC AAC GAA CGA TTC ATG ACC GTG
Val Thr Trp Thr Thr Asp Gln Leu Gly Lys Ala Phe Phe Tyr His Asp Arg Ile Leu Asn Glu Arg Phe Met Thr Val

CTG ACC TGA TGGTCCGG
Leu Thr ***
```

FIG. 3. Complete nucleotide sequence of the *xylE* gene and the corresponding amino acid sequence. The sequence shown starts 100 base pairs 3' to the leftmost *Hpa* II site in Fig. 4 and ends at the *Hpa* II site on the extreme right of Fig. 4. The putative ribosome binding site is underlined. The coding region begins at base 1 and ends at the nonsense codon (★★★).

the time of its appearance after spraying colonies is directly related to the specific activity of the enzyme. The subjective nature of such interpretations, however, must be viewed with caution. In any event, the specific activity of the enzyme produced is rapidly confirmed by a spectrophotometric assay (24).

Expression of *xylE* was directed by fragments of chromosomal DNA originating from *B. subtilis*, *B. pumilus*, *B. licheniformis*, and *E. coli* (Table 1). Fragments were inserted into the unique *Bam*HI site of pTG402 and transformed into *B. subtilis* protoplasts, and cells that expressed *xylE* were selected directly on protoplast regeneration medium (19). Because the *Bam*HI site lies approximately 800 nucleotide base pairs from the ATG initiation codon of *xylE*, we assume that transcriptional termination signals are absent in this region. The interpretation is consistent with evidence that *xylE* is the second gene of a regulated operon on the *Pseudomonas* TOL plasmid (22, 23).

The nucleotide sequence of the ribosome binding site nine bases from the translational initiation codon is of considerable interest. The relatively strong complementarity between the sequence and that of the 3′ end of *B. subtilis* 16S rRNA (1, 37) may provide the stable interaction required for the initiation of translation by ribosomes of Gram-positive bacteria (1). The pyrimidine residue (cytidine) 5′ to the initiation codon and the adenosine residue 3′ to the ATG may also help to stabilize the translational initiation complex (39, 40). The sequence data, together with supporting evidence from polyacrylamide gel electrophoresis (Fig. 2) of the translation product, show that translation of *xylE* mRNA in *B. subtilis* most likely commences at the same initiation codon as that in *Pseudomonas* and *E. coli*.

Plasmid vectors suitable for cloning fragments of DNA that carry transcriptional promoter or termination signals or translational initiation signals have been described elsewhere (2–8). The system described here offers several advantages. (*i*) Specially prepared indicator plates are not necessary to monitor *xylE* gene expression. (*ii*) The $CatO_2$ase substrate, catechol, is very inexpensive as compared with the chromogenic β-galactosidase substrate, 5-bromo-4-chloro-3-indolyl β-D-galactoside, needed to measure very low levels of *lacZ* gene expression in *E. coli* (4). (*iii*) As our results indicate (Table 1), the catechol spray test is very sensitive in that it detects colonies of cells that produce low to quite appreciable levels of $CatO_2$ase. Coloration is observed only for those cells that express *xylE*; background coloration is not observed for cells in which *xylE* is present but remains unexpressed. (*iv*) $CatO_2$ase specific activities are determined by a simple spectrophotometric assay.

The unique *Bam*HI, *Hpa* I, and *Kpn* I sites on the cloning vector pTG402 represent potential cleavage sites for the insertion of DNA fragments that carry signals to promote gene expression in *B. subtilis*. Plasmids that express *xylE* might also be used to isolate transcriptional termination signals or as cloning vectors for the construction of other recombinant plasmids. Insertion of foreign DNA into the *xylE* structural gene would lead to rapid identification of recombinant plasmids because colonies of host cells would not produce $CatO_2$ase and thus would not change color with catechol.

The chromogenic assay we have constructed offers an approach to the development of plasmid gene expression vectors. Translation of *xylE* mRNA in *Pseudomonas*, *E. coli*, and *B. subtilis* indicates that our system may be applicable to a wide range of host organisms.

We thank Professors P. Chambon and P. Kourilsky for their continued interest in our work and for their critical review of the manuscript. We gratefully acknowledge K. Timmis for his advice on the initial work with *Pseudomonas* and P. Sautière for determining the NH_2-terminal amino acid sequence. We thank R. Lathe, M. Jaye, and M. P. Kieny for their constructive suggestions and I. Batra for her expert secretarial assistance.

1. McLaughlin, J. R., Murray, C. L. & Rabinowitz, J. C. (1981) *J. Biol. Chem.* **256**, 11283–11291.
2. West, R. W., Neve, R. L. & Rodriguez, R. L. (1979) *Gene* **7**, 271–288.
3. An, G. & Friesen, J. D. (1979) *J. Bacteriol.* **140**, 400–407.
4. Casadaban, M. J. & Cohen, S. N. (1980) *J. Mol. Biol.* **138**, 179–207.
5. Casadaban, M. J., Chou, J. & Cohen, S. N. (1980) *J. Bacteriol.* **143**, 971–980.
6. Gentz, R., Langner, A., Chang, A. C. Y., Cohen, S. N. & Bujard, H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4936–4940.
7. Williams, D. M., Duvall, E. J. & Lovett, P. S. (1981) *J. Bacteriol.* **146**, 1162–1165.
8. Goldfarb, D. S., Doi, R. H. & Rodriguez, R. L. (1981) *Nature (London)* **293**, 309–311.
9. Worsey, M. J. & Williams, P. A. (1975) *J. Bacteriol.* **124**, 7–13.
10. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 47–55.
11. Spizizen, J. (1958) *Proc. Natl. Acad. Sci. USA* **44**, 1072–1078.
12. Primrose, S. B. & Erhlich, S. D. (1981) *Plasmid* **6**, 193–201.
13. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523.
14. Guerry, P., Le Blanc, D. J. & Falkow, S. (1973) *J. Bacteriol.* **116**, 1064–1066.
15. Radloff, R. W., Bauer, W. & Vinograd, J. (1967) *Proc. Natl. Acad. Sci. USA* **57**, 1514–1521.
16. Morrison, D. A. (1977) *J. Bacteriol.* **132**, 349–351.
17. Lederberg, E. M. & Cohen, S. N. (1974) *J. Bacteriol.* **119**, 1072–1074.
18. Boylan, R. J., Mendelson, N. H., Brooks, D. & Young, F. E. (1972) *J. Bacteriol.* **110**, 281–290.
19. Chang, S. & Cohen, S. N. (1979) *Mol. Gen. Genet.* **168**, 111–115.
20. Marmur, J. (1961) *J. Mol. Biol.* **3**, 208–218.
21. Nakazawa, T., Inouye, S. & Nakazawa, A. (1980) *J. Bacteriol.* **144**, 222–231.
22. Inouye, S., Nakazawa, A. & Nakazawa, T. (1981) *J. Bacteriol.* **145**, 1137–1143.
23. Franklin, F. C. H., Bagdasarian, M., Bagdasarian, M. M. & Timmis, K. N. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7458–7462.
24. Sala-Trepat, J. M. & Evans, W. C. (1971) *Eur. J. Biochem.* **20**, 400–413.
25. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265–275.
26. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9**, 309–321.
27. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol.* **143**, 161–178.
28. Downing, R. & Broda, P. (1979) *Mol. Gen. Genet.* **168**, 189–191.
29. Bolivar, F., Rodriguez, R., Betlach, M. & Boyer, H. (1977) *Gene* **2**, 75–93.
30. Sutcliffe, J. G. (1979) *Cold Spring Harbor Symp. Quant. Biol.* **43**, 77–90.
31. Canosi, U., Morelli, G. & Trautner, T. A. (1978) *Mol. Gen. Genet.* **166**, 259–267.
32. Tanaka, T. (1979) *Mol. Gen. Genet.* **175**, 235–237.
33. Davis, B. (1964) *Ann. N.Y. Acad. Sci.* **121**, 404–427.
34. Nozaki, M., Ono, K., Nakazawa, T., Kotani, S. & Hayaishi, O. (1968) *J. Biol. Chem.* **243**, 2682–2690.
35. Nozaki, M. (1979) *Topics Curr. Chem.* **78**, 145–186.
36. Shine, J. & Dalgarno, L. (1975) *Nature (London)* **254**, 34–38.
37. Moran, C. P., Lang, N., Le Grice, S. F. J., Lee, G., Stephens, M., Sonenshein, A. L., Pero, J. & Losick, R. (1982) *Mol. Gen. Genet.* **186**, 339–346.
38. Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246**, 40–41.
39. Taniguchi, T. & Weissman, C. (1978) *J. Mol. Biol.* **118**, 533–565.
40. Ganoza, M. C., Fraser, A. R. & Neilson, T. (1978) *Biochemistry* **17**, 2769–2775.