lacZ- and aph-Based Reporter Vectors for In Vivo Expression Technology

BAEK, CHANG-HO AND KUN-SOO KIM*

Department of Life Science, Sogang University, Sinsoo-Dong 1, Mapo-Gu, Seoul 121-742, Korea

Received: May 6, 2003
Accepted: June 16, 2003

Abstract Three vectors, pSG1, 2, and 3, which facilitate in vivo expression technology (IVET) in Gram-negative bacteria, were developed. Vectors pSG1 and 2 are derivatives of ColE1, and pSG3 is a derivative of an R6K replicon. These vectors contain oriT sites that allow mobilization when the RK2 Tra functions are provided in trans. These vectors contain promoterless lacZ (pl-lacZ) and promoterless aph (pl-aph) transcriptionally fused together, which allow qualitative and quantitative measurements of the expression of genes in the genome of bacterial cells. pSG1 and 3 contain gentamicin-resistance genes, and pSG2 carries a streptomycin-/spectinomycin-resistance gene, allowing for selection of recombinants generated by a single crossover between a library fragment cloned into a pSG vector and the identical region in the genome of a bacterial species from which the library fragment originated. These vectors were successfully applied to the generation of random fusions at high rates in the genomes of four representative Gram-negative bacteria. In addition, the expression level of β-galactosidase and the degree of resistance to kanamycin in cells with fusions generated by these vectors were found to be linearly correlated, proving that these vectors can be used for IVET.

Key words: Reporter vector, IVET, Rhodobacter sphaeroides, Agrobacterium tumefaciens, Vibrio vulnificus

The identification of genes expressed under certain conditions and the quantitative measurement of the expression levels of these genes are crucial to understanding the nature of an organism. Especially for pathogenic bacteria, such studies have provided valuable insight into the molecular mechanisms underlying interactions with host organisms. In vivo expression technology (IVET) has served as a powerful method for the identification of virulence genes from pathogenic microorganisms. This technique is based on the assumption that genes associated with virulence are expressed at certain points of time during infection at a higher level than in vitro.

The development of a properly designed vector is one of the important elements for the successful application of IVET. The IVET system is based on a suicide vector, which contains reporter genes that lack their own promoters, and a genetic marker for selecting cells with the reporters inserted into their genome. The general scheme of IVET exemplified by pSG vectors is illustrated in Fig. 1. Genomic library fragments from a studied microorganism are cloned into the region immediately upstream from the reporter genes in the vector, so that the cloned fragment is transcriptionally fused to the reporter genes (Fig. 1a). The resulting constructs are introduced into the cells of the microorganism from which the library fragment originated. By selecting cells showing resistance to the antibiotic conferred by the vector, we can obtain cells carrying fusions of reporters into the genome generated by homologous recombination (Fig. 1b). In general, IVET vectors carry two reporters: one for positive selection of cells carrying fusions of genes, which are expressed in test conditions (e.g. host milieu), and the other for monitoring the expression of the gene in vitro (Fig. 1c). Generally, those genes that are expressed in hosts but not in vitro are considered to be associated with survival, propagation, and virulence in the test conditions. In IVET, a single crossover event results in the generation of an extra wild-type copy of the gene inserted by fusion (Fig. 1b). Consequently, IVET is advantageous over other fusion techniques such as transposon-based reporter fusions, which abolish the function of the inserted genes and, furthermore, could exert a polar effect on downstream genes. Such insertions could result in biased cell physiology and even abolish the viability of cells if the inserted gene is essential. Therefore, IVET has been widely applied to the identification of bacterial genes that are expressed at a higher level in specific environmental conditions, especially in host environments, during infection by pathogenic bacteria. Since IVET was successfully employed for the isolation of novel virulence factors in the plant pathogen Xanthomonas...
We describe herein the development of a series of new IVET vectors, pSG1, 2, and 3, which contain a promoterless *aph* (a kanamycin-resistance gene) for qualitative selection and a promoterless *lacZ* for quantitative measurements of the expression of genes fused to the reporters on the vectors. One of the vectors, pSG3, carries an R6K replicon and, therefore, this vector can also be used in *Escherichia coli*, in which R6K derivatives cannot exist as episomes, unless the cells produce the \( \pi \) factor. These pSG vectors successfully generated random fusions of the reporters in the genome of four representative Gram-negative bacteria; a general laboratory *E. coli* strain, a human pathogen *Vibrio vulnificus*, a plant pathogen *Agrobacterium tumefaciens*, and a free-living species *Rhodobacter sphaeroides*.

**Fig. 1.** Selection and retrieval of genes, the expression of which is induced in specific environmental conditions, using pSG vectors. 
(a) Sau3AI-partial digested fragments of the genomic DNA from the bacterial species of interest are cloned into the BamHI site of pSG to construct the genomic library. Using an *E. coli* strain providing Tra in trans, the library is mobilized into the bacterium, from which the library clones were originated. Exconjugants are selected in the medium containing an antibiotic, to which resistance is encoded by the gene (\( \text{AB} \)) in the vector. In the exconjugants, pSG cannot replicate, and hence the vector was inserted into the genome by a homologous recombination event. (b) These cells are exposed to the specific conditions of interest, and cells resistant to kanamycin are selected. If reporter genes in pSG are properly fused to a gene that is induced in these conditions, those cells become resistant to kanamycin. The expression of the inserted gene is quantitatively assessed by \( \beta \)-galactosidase activity. (c) Genomic DNA is purified from the cells and digested with the restriction enzyme that cuts the plasmid at one position between the insert and \( \text{bla} \) of the vectors. The digested DNA fragments are religated, and the reactions are transformed into an *E. coli* strain that allows the replication of the vector. (d) Transformants resistant to ampicillin should carry a plasmid that contains a host DNA fragment of the region upstream to the reporters of pSG. The DNA nucleotide sequences of the gene are determined using a DNA oligomer as a primer (denoted as a blank arrow) with nucleotide sequence complementary to the 5' end of pl-aph in the case of pSG3 or of pl-lacZ in the cases of pSG1 and pSG2. Nucleotide sequences of the primers used are 5'-TAAATGACATCCATGGTG-3' and 5'-TTCCAGTCAGACGCTTGG-3', respectively.

*campestris* [14, 23], and the human pathogen *Salmonella typhimurium* [20, 21], the validity of this method has been confirmed in various microorganisms, including Gram-negative bacteria (*e.g.* *Actinobacillus pleuropneumoniae* [31], *Porphyromonas gingivalis* [40], *Pseudomonas aeruginosa* [39], *P. fluorescens* [26], *P. putida* [17], *P. syringae* [2], *Shigella flexneri* [1], *Vibrio cholerae* [3], *Yersinia enterocolitica* [41]), Gram-positive bacteria (*e.g.* *Staphylococcus aureus* [19], *Streptococcus gordonii* [12], *Listeria monocytogenes* [7]), and even fungi (*e.g.* *Candida albicans* [15, 30], *Histoplasma capsulatum* [27]).

**MATERIALS AND METHODS**

**Bacterial Strains and Culture Conditions**

*E. coli* S17-1 [32] was used for the amplification, maintenance, and conjugal transfer of derivatives of pSG1 and 2, and *E. coli* S17-1:3apir [5] was used for derivatives of pSG3. *A. tumefaciens* C58 [9, 30] and *R. sphaeroides* 2.4.1 [35] are naturally-occurring *lac* strains, and *E. coli* CSH26 is a *lacZ* derivative of *E. coli* K-12 [22]. The *V. vulnificus* strain MO6ΔZ is a *lacZ*-null mutant derived from a wild-type strain MO6-24/0 [13]. As rich media, LB (Difco, Detroit, MI, USA) for *E. coli*, *R. sphaeroides* [11], and *V. vulnificus* [13], and NB (Difco, Detroit, MI, U.S.A.) for *A. tumefaciens* [10] were used. *E. coli* strains were grown at 37°C, and *A. tumefaciens*, *R. sphaeroides*, and *V. vulnificus* were grown at 28°C. For *E. coli*, ampicillin (Ap) at 50 µg/ml, gentamicin (Gm) at 15 µg/ml, and kanamycin (Km) at 25 µg/ml were used. For *V. vulnificus*, ampicillin (Ap) at 100 µg/ml, gentamicin (Gm) at 30 µg/ml, and kanamycin (Km) at 150 µg/ml were used. For *R. sphaeroides*, spectinomycin (Sm) at 25 µg/ml, spectinomycin (Sp) at 25 µg/ml, and Km at 5 µg/ml were used.

**Chemicals**

All of the chemicals and antibiotics used in this study were purchased from Sigma Chemical Company (St. Louis, MO, U.S.A.).

**DNA Manipulations and Transformation**

Standard recombinant DNA techniques were used as described by Sambrook and Russell [29]. Restriction
enzymes, T4 DNA ligase, and T4 DNA polymerase were purchased from Promega (Madison, WI, U.S.A.). ExoIII and S1 nuclease were purchased from MBI Fermentas (Hanover, MD, U.S.A.).

Construction of Genomic Libraries of A. tumefaciens, R. sphaeroides, V. vulnificus, and E. coli in pSG Vectors

Approximately 100 µg of genomic DNA was partially digested with Sma3AI, and the DNA fragments were size-fractionated in a 5–40% (w/v) sucrose gradient by centrifugation at 26,000 rpm at 20°C for 24 h in a Beckman SW41 rotor. Fractions containing DNA fragments of 3–5 kb were pooled and precipitated by cold 95% (v/v) ethanol. The purified DNA fragments were ligated into the unique BamHI site of the pSG vectors. The ligated DNAs were introduced by electroporation (at 2.5 kV, 25 µF, and 200 Ω using 0.2 cm-cuvettes) into E. coli S17-1 for derivatives of pSG1 and 2, and E. coli S17-1::Ap′ for derivatives of pSG3 [4]. The colonies that formed on Luria agar medium containing appropriate antibiotics were pooled and stored. Transformants were used to introduce the pSG derivatives into cells by conjugation [32, 37].

Measurement of β-Galactosidase Activity

β-Galactosidase activity was quantitatively measured with o-nitrophenyl-β-D-galactopyranoside (ONPG) as a substrate by the method described previously [22, 24].

RESULTS AND DISCUSSION

Construction of pSG1

An 854-bp fragment containing pl-aph from pUC4K [36] was amplified by PCR using two primers with the DNA sequences of 5'-TACGGATCCAGTAAATACGAGGGTGTATG-3' and 5'-ACCAGATCTGATTGAAAACCT-CATCG-3'. The underlined bases denote bases to generate BamHI and BglII restriction sites, respectively. The PCR product was digested with BamHI and BglII, and ligated into the BamHI site of pBluescript SK(−) (Strategene, La Jolla, CA, U.S.A.) to construct pBL-Km. This plasmid was digested with BamHI and PstI, and the sticky ends were blunted with T4 DNA polymerase. This product was self-ligated to generate pBL-KmdBP. A 1.0-kb EagI fragment containing the gentamicin-resistance cassette from pFastBacHl (Gibco BRL, Grand Island, NY, U.S.A.) was cloned into the Eagl site of pBL-KmdBP to construct pBLKmdBPGm. Plasmid pRSET-A (Invitrogen, Carlsbad, CA, U.S.A.) was digested with NdeI and BamHI, treated with T4 DNA polymerase to produce blunt ends, and religated to generate pRSET-dNB. A 2.5-kb SaI fragment from pRK415 [10], containing the oriT region, was cloned into the SaI site of pUC19. From the resulting construct, the 2.5-kb HindIII fragment containing the oriT region was cloned into the HindIII site of pRSET-dNB. The resulting plasmid, named pRSET-dNBt, was ligated with the 1.9-kb EcoRI-SacI fragment containing the gentamicin-resistance cassette and promoterless-aph gene from pBL-KmdBPGm to construct pRSET-TGK. This plasmid was digested with PstI and SalI, and treated with EcoHI and S1 nuclease, followed by a Klenow treatment to remove the approximately 1.5-kb region between the SpI site and oriT region. The resulting DNA fragment was self-ligated to construct pTGKND28. A 4.6-kb BamHI-EcoRI fragment containing the promoterless-lacZ (pl-lacZ) and the Km′ cassette from mini-Tn5lacZ [6] was cloned into the pBluescript SK(−). The resulting plasmid was digested with NotI to remove the 1.5-kb fragment containing the Km′ cassette, and the resulting 6.0-kb fragment containing the pl-lacZ was self-ligated to construct pBS-pl-lacZ. From this plasmid, a 3.1-kb SphI-EcoRI fragment containing the pl-lacZ was cloned into the SpHl and EcoRI sites of pTGKND28 to generate pTGKZ. The 2.0-kb BamHI-SacI fragment containing an N-terminal two-thirds portion of pl-lacZ of pTGKZ was replaced with a 1.9-kb BamHI-SacI fragment containing the N-terminal two-thirds portion of pl-lacZ of the promoter selection vector pRG970 [38], to construct pSG1 (Fig. 2).

Construction of pSG2

The vector pSG1 was partially digested with XbaI, and a 7.9-kb fragment, in which a 0.9-kb XbaI fragment containing the gentamicin-resistance gene was removed, was self-ligated to construct pSG523ΔGm. A 2.0-kb HindIII fragment of the Ω cassette containing the Sm'/Sp′ gene from pHP45Ω-Sm/Sp [25] was cloned into pBluescript SK(−) to construct pBS-ΩSm/Sp, and a 2.0-kb PstI-SalI fragment from the plasmid was cloned into pSG523ΔGm to construct pSG621. This plasmid was partially digested with XbaI, and the XbaI site at the region downstream to pl-aph was removed by filling in the sticky ends using T4 DNA ligase and self-ligation to construct pSG2 (Fig. 3).

Construction of pSG3

Three BamHI sites in pKAS32 [33] were removed by digesting with BamHI followed by filling in the sticky ends with T4 DNA polymerase and religation. The 715-bp rpsL region was removed by digesting with SalI and self-ligation. A 3.1-kb SalI-NotI fragment containing pl-lacZ from pSG1 was cloned into pKASdB3 to construct pBSG4a. A 1469-bp DNA region in pBRR1MC5 [15, 28] containing the MCS and gentamicin-resistance gene was amplified by PCR using the following primers; forward (T7 primer), 5'-GTAATACGACTCACTATAGGGT-3'; reverse primer, 5'-CATCTGAAATCCGCCCGACAC-3'. The underlined bases represent an artificially introduced XbaI site. The PCR product was digested with XbaI and cloned into the XbaI site of pBSG4a to construct pBSG4Gm. The 854-bp
**Fig. 2.** Construction of pSG1.

The IVET vector pSG1 is a derivative of pBS-pl-λacZ, in which pl-λacZ is transcriptionally fused with pl-aph, and contains gentamicin-resistance (Gm') and ampicillin-resistance (bla) genes. Abbreviations; pl-aph, promoterless kanamycin-resistant gene; pl-λacZ, promoterless lacZ; bla, ampicillin-resistant gene; Gm, gentamicin-resistant gene; Ω-Sm/Sp, Sm-Sp-resistant gene in Ω cartridge which contains translation/transcription stop sequences; A, ApaI; Bm, BamHI; Bg, BglII; E1, EcoRI; E5, EcoRV; H, HindIII; K, KpnI; Nd; No, NotI; P, PstI; S1, SacI; Sa, SalI; Sf, SfiI; Sm, SmaI; Sph, SphI; Xb, XbaI; Xh, XhoI. Restriction sites that can be used for the retrieval of the chromosomally integrated pSG and flanking genomic DNA (Fig. 1d) are denoted by asterisks (*).

PCR product containing pl-aph described above was digested with BamHI and BglII and inserted into the BamHI site of pBSG4Gm to obtain pSG3 (Fig. 4).

**Test of pSG Plasmids for Feasibility as IVET Vectors**

If a vector can replicate or exist as an epimer in host cells, the vector cannot be employed for IVET. We tested to see whether or not library clones in the three pSG vectors could replicate or exist as epimeres in four representative Gram-negative bacterial strains, *E. coli* CSH26, *A. tumefaciens* C58, *R. sphaeroides* 2.4.1, and *V. vulnificus* MO6AZ by plasmid DNA preparations and Southern hybridization analysis. In *R. sphaeroides* and *A. tumefaciens*, derivatives of pSG1 and pSG2 containing library clones of
genomic elements from the corresponding parental cells could not autonomously replicate, and were maintained after being integrated into the genome of the bacteria (data not shown). However, in *E. coli* and *V. vulnificus*, the library clones could exist as episomes, when cells were grown under antibiotic selection (data not shown). In contrast, derivatives of pSG3 can neither replicate nor exist as episomes in any of the four strains and could be stably inserted in the genome (data not shown). These results indicated that pSG1 and 2, which are ColE1 derivatives, can be maintained in *V. vulnificus* MO6ΔZ, and pSG3 can be employed for IVET in this species, which apparently does not provide a π-factor. This was verified by plasmid preparations from the strain (data not shown). In addition, we could not detect any integrations of pSG vectors onto their host genomes by non-ε homologous recombination, as confirmed by Southern hybridization analysis (data not shown).

**Fig. 3.** Construction of pSG2.

pSG-2 is a derivative of ColE1, and contains multiple cloning sites (MCS) located between the Ω-Sm/Sp cassette and pl-locZ. The same abbreviations are used as in the legend to Fig. 2.
pSG vectors contain promoterless lacZ and aph genes transcriptionally fused together, which allow qualitative and quantitative measurements of the expression of the genes in the genomes of bacteria. We tested the level of the expression of these two reporter genes from cells carrying the plasmids randomly integrated in the genomes of four bacterial species. As shown in Fig. 5, the degree of cell growth on solid media containing kanamycin was closely related with the level of the expression of lacZ, as measured by β-galactosidase activity in fusions generated by any of the three pSG vectors. These results suggest that the pSG vectors are all suitable for IVET.

In summary, pSG vectors were designed to contain genes conferring resistance to four antibiotics [ampicillin (pSG1, 2, and 3), gentamicin (pSG1 and 3), and streptomycin and spectinomycin (pSG2)], providing a broad choice of useful markers to be used for the selection of cells carrying fusions generated by the vectors. These vectors all have unique BamHI sites immediately upstream from the reporter genes, which can be used to clone genomic library fragments generated by Sau3AI. These vectors also carry
Fig. 5. Expression of β-galactosidase and resistance to kanamycin of cells with fusions generated by pSG vectors are linearly correlated.
A. Growth of cells on solid media containing kanamycin. Cells with fusions were randomly selected and grown on appropriate minimal medium containing kanamycin as described in Materials and Methods. Growth was visually assessed after two days. Numerical denotations represent the numbers of randomly selected clones. B. Quantitative measurement of β-galactosidase activities from the same sets of fusion clones tested in A. a, A. tumefaciens C58 with pSG1-fusions; b, R. sphaeroides 2.4.1 with pSG2-fusions; c, V. vulnificus MO6ΔZ with pSG3-fusions; d, E. coli CSH26 with pSG3-fusions.
oriT, hence, facilitate a convenient introduction to the desired recipient cells using a donor strain such as S17-1 providing Tra in trans. The genes inserted by these reporter vectors can be retrieved using the unique restriction enzyme site present in these vectors (Fig. 1d). pSG3 is particularly practical for IVET in E. coli or bacterial species such as V. vulnificus, in which derivatives of CoIE1 can exist episomally. We are currently carrying out IVET on A. tumefaciens and V. vulnificus using the pSG vectors to identify novel genes associated with interactions with host organisms (unpublished results). These vectors will also be useful in other related species, and provide useful molecular genetic tools to identify novel genes playing significant physiological roles in various environmental conditions and to monitor the expression of those genes.

Acknowledgments

This work was supported by the 21C Frontier Microbial Genomics and Applications Center Program, Ministry of Science & Technology (Grant MG02-0201-004-1-0-0), Republic of Korea and also partly by a grant No. 2000-2- of Science & Technology Genomics and Applications Center Program, Ministry , hence, facilitate a convenient introduction to the site present in these vectors (Fig. 1d). pSG3 is particularly practical for IVET in E. coli or bacterial species such as V. vulnificus, in which derivatives of CoIE1 can exist episomally. We are currently carrying out IVET on A. tumefaciens and V. vulnificus using the pSG vectors to identify novel genes associated with interactions with host organisms (unpublished results). These vectors will also be useful in other related species, and provide useful molecular genetic tools to identify novel genes playing significant physiological roles in various environmental conditions and to monitor the expression of those genes.

REFERENCES


