Center for Eukaryotic Structural Genomics
Protein Structure Initiative

USE OF *BACILLUS SUBTILIS sacB* (LEVANSUCRASE) AS A NEGATIVE SELECTABLE MARKER IN FLEXI®VECTOR CLONING

Russell L. Wrobel, Zachery Kunkel, and Brian G. Fox
University of Wisconsin-Madison, Department of Biochemistry, 433 Babcock Drive, Madison, WI USA 53706-1549, http://www.uwstructuralgenomics.org

Abstract

Flexi®Vector system is a directional restriction/ligation cloning method developed by Promega (Madison, WI). In the plasmids used for this cloning method, a lethal barnase gene insert is flanked by two rare-cutting restriction enzymes, *Sgf* I and *Pme* I. During the cloning step, the barnase insert is removed by restriction digestion and is replaced by the similarly digested, desired coding sequence. Replacement of the barnase gene thus permits growth of the successfully transformed host strain, while preventing growth of transformants where the barnase insert is still present. Experiments using plasmids containing the barnase gene must be propagated in an *Escherichia coli* strain over expressing the inhibitor barstar (e.g., Promega BR610); importantly these experiments include preparation of reagent plasmids and new vector engineering. Arising from high toxicity of the barnase selection, the use of the BR610 strain can be less robust than desired. In order to improve on this cloning system, we investigated the use of the *sacB* gene as an alternative selection method. Expression from *sacB* confers sensitivity to sucrose in a wide variety of gram positive and negative bacteria and thus has been used extensively for the last twenty years as a negative selectable marker.[1,2,3] Replacement of the barnase gene with the *sacB* gene gave a modified Flexi®Vector that performed in cloning experiments with efficiency comparable to the previous barnase selection. With this modification, Flexi®Vector plasmids containing the *sacB* gene can be transformed into any *E. coli* strain, which greatly simplifies the process of vector engineering. For plasmid reagent preparation, propagation in a *dam*⁻ strain is still required in order to provide the methylation pattern required for efficient ligation at the *SgfI* site during the cloning step. INV110 is a highly competent *dam*⁻ strain that can be purchased through Invitrogen (Carlsbad, CA). Improved versions of CESG expression vectors containing the *sacB* insert, including some for *E. coli*, wheat germ cell-free translation, and insect cells will be available from the Protein Structure Initiative Materials Repository, website: http://www.hip.harvard.edu/PSIMR/

An Overview of Cloning into Flexi®Vector Plasmids[6]

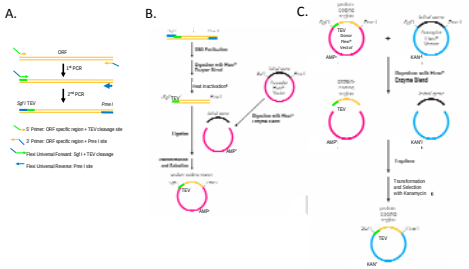

Figure 1A. Two step PCR to append a TEV site and restriction sites. ORF specific PCR primers are designed to append part of the Tobacco Etch Virus (TEV) protease cleavage site to the 5' primer and the *Pme* I site to the 3' primer. The universal primers for the second PCR step completes the TEV site and adds the *Sgf* I site. The PCR primers also append the TEV site to the 5' end so the amino acids encoded in part by the *Sgf* I site can be removed with any amino terminal tags that may be designed in the vector.

Figure 1B. Cloning a protein-coding region into Flexi®Vector plasmids. After amplification, the PCR product is purified to remove the DNA polymerase and primers and is digested with *Sgf* I and *Pme* I. The heat-inactivated, digested PCR product is ligated into an acceptor Flexi®Vector that has been digested with *Sgf* I and *Pme* I. Following transformation, the cells are selected with the appropriate antibiotic for the particular Flexi®Vector used. Cells that are transformed with non-recombinant acceptor plasmid do not grow due to the presence of the lethal gene. In the Promega Flexi®Vector plasmids the lethal gene is the RNase, Barnase. The adenosine of the *Sgf* I recognition site (GCGATCGC) must be unmethylated for efficient ligation. Therefore, the acceptor plasmid must be propagated in strains that lack *dam* methylation.

Figure 1C. Transfer of a protein-coding region between Flexi®Vector plasmids. The donor Flexi®Vector containing the protein-coding region is mixed with an acceptor Flexi®Vector that has a different antibiotic resistance. The two plasmids are digested with *Sgf* I and *Pme* I, and the mixture is ligated and transformed into *E. coli*. The cells are plated on the appropriate selective medium for the acceptor Flexi®Vector.

B. subtilis sacB Gene

Figure 2. Annotated DNA sequence of the *sacB* region.[4] The green arrows show the -10 and -35 region of the *sacB* promoter. This is followed by a small ORF that may act as a transcriptional attenuator included in the construct to maintain the proper context of the gene. The *sacB* coding region is 1422 bp in length and encodes levansucrase. A fairly large inverted repeat lies a the 3' end of the *sacB* region. This is thought to act as a transcriptional terminator.

Levansucrase is a secreted enzyme that catalyzes the hydrolysis of sucrose to glucose and fructose. The fructosyl units are then polymerized to form levan. The expression of *sacB* in *E. coli* in the presence of sucrose is lethal. The mechanism behind this lethality is not fully understood.

Flexi®Vector Plasmids for Expression in *E. coli*

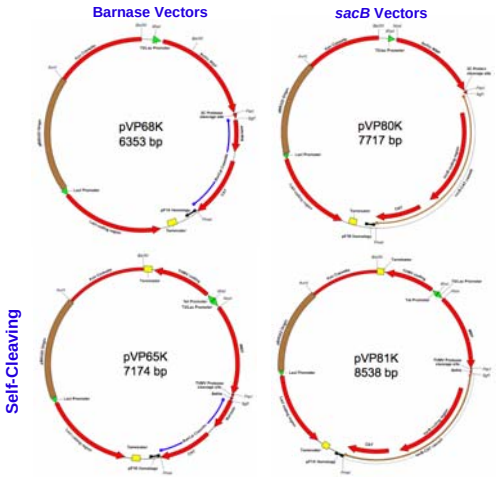

Figure 3. Maps of Flexi®Vector plasmids for protein expression in *E. coli*. On the left our production vectors that contain the lethal gene barnase together with the chloramphenicol acetyltransferase (CAT) gene bordered by the cloning sites *Sgf* I and *Pme* I. Versions of these vectors containing *sacB* (right side) were made by swapping out the barnase gene for the *sacB* gene (shown in **Figure 2**). The *sacB* gene was obtain by PCR amplification from *B. subtilis* genomic DNA. The pVP68K and pVP80K vectors (top) will produce proteins fused to an amino-terminal His_6-MBP tag. The TEV protease restriction site is included in the PCR amplified product so that the tags can be cleaved off along with the AIA amino acid residues that are encoded by the *Sgf* I site (see **Figure 1A**).

The pVP65K and pVP81K vectors (bottom) also fuse MBP to the amino end of the target protein. However, the MBP fusion protein is proteolysed *in vivo* by the Tobacco Vein Mottling Virus (TVMV) protease expressed from the same plasmid because the TVMV protease site is present in the linker region between the MBP and the *Sgf* I site.

These vectors are modular in design. Several important regions, such as the antibiotic cassette, promoter, solubility tag (MBP), and linker region are bordered by unique restriction sites and can easily be substituted with other DNA sequences. The *sacB* vectors make this vector engineering easier since the plasmid variants do not have to be transformed and propagated in the BR610 strain. For more details of the expression of proteins from these plasmids see the poster and presentation by Dave Aceti entitled, "A Control Workgroup of 24 Protein Targets Used to Test Platform Changes."

Flexi®Vector Plasmids for Cell-Free Expression

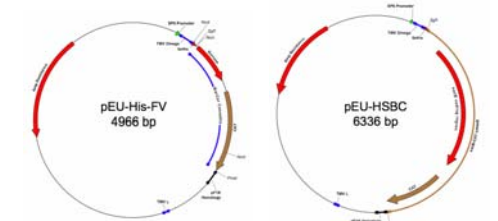

Figure 4. Maps of Flexi®Vector plasmids for cell-free protein production. The *sacB*-containing pEU–HSBC vector was derived from pEU-His-FV by swapping the barnase gene for the *sacB* gene. The SP6 promoter is used for the *in vitro* production of mRNA. The TMV Omega fragment acts as a translational enhancer in the wheat germ translation. The region labeled pF1K homology is present in all of our Flexi®Vector plasmids and is needed for efficient transfer of the target ORF between vectors. If there is no homology between vectors, it is possible for the two vectors to ligate together during the transfer reactions to yield a plasmid with resistance to both antibiotics and having no lethal gene. For the last several years we have been testing all of our targets in both cell free and *E. coli* platforms. Our process is to first clone the ORF into a cell-free expression vector, verify the sequence, then transfer the cloned ORF to an *E. coli* expression vector.

sacB Confers Sucrose Sensitivity in *E. coli*

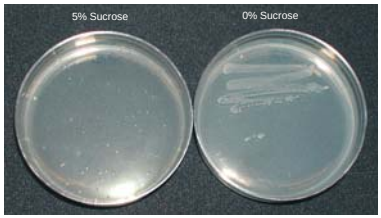

Figure 5. *E. coli* harboring the pVP80K plasmid was grown overnight in liquid medium containing kanamycin and chloramphenicol. The culture was then streaked onto YT plates containing the two antibiotics with 5% (w/v) sucrose or without added sucrose (0% sucrose). Growth is only observed on the plate containing no sucrose. The dots on the plate containing 5% sucrose are bubbles under the agar.

The *sacB* Gene Can Be Used for the Lethal Gene in Flexi®Vector Cloning

Target in the Control Work Group	Number of Colonies Observed	Colony PCR Positives Out of Two Screened
A1	17	2
A3	27	2
A5	45	2
A6	32	2
A7	4	2
A9	24	0
A10	39	2
A11	27	2
A12	51	2
B1	14	2
B2	6	2
B3	47	1
B4	32	2
B5	22	2
B6	29	2
B7	53	1
B8	6	1
B9	10	2
B10	25	2
B11	44	2
B12	53	2

Table 1. Flexi®Vector transfer of ORFs from pVP68K to pEU-HSBC. Twenty-one ORFs from the CESG control work group in pVP68K were transferred to the *sacB*-containing vector, pEU-HSBC, and plated on plates containing ampicillin and 5% sucrose. The first column shows the number of colonies for each transformation. The second column shows how many out of two colonies screened per ORF were positive in a PCR screen. Most colonies screened positive. This cloning effiency is comparable to the to the effiency we observed with barnase selection.[5] We were never able to get a positive colony for A9 even after multiple attempts, suggesting that there may have been a handling error with that sample.

References

(1) Gay P, Le Coq D, Steinmetz M, Berkelman T, and C Kado. (1985) Positive selection procedure for entrapment of insertion sequence elements in gram negative bacteria. *J Bacteriol* 164, 918–921.

(2) Pierce JC, Sauer B, and N Sternberg.(1992) A positive selection vector for cloning high molecular weight DNA by the bacteriophage P1 system: improved cloning efficacy. *Proc Natl Acad Sci USA* 89, 2056–2060.

(3) Frengen E, Weichenhan D, Zhao B, K. Osoegawa K, van Geel M, and PJ de Jong. (1999) A modular, positive selection bacterial artificial chromosome vector with multiple cloning sites. *Genomics* 58, 250–253.

(4) Steinmetz M, Le Coq D, Aymerich S, Gonzy-Treboul G, and P Gay. (1985) The DNA sequence of the gene for the secreted *Bacillus subtilis* enzyme levansucrase and its genetic control sites. *Mol Gen Genet* 200, 220-228.

(5) Blommel PG, Martin PA, Wrobel RL, Steffen E, and Fox BG. (2006) High-efficiency single-step production of expression plasmids from cDNA clones using the Flexi®Vector cloning system. *Prot Exp Purif* 47, 562-570.

(6) Promega Corporation. (2005) Flexi®Vector Systems Technical Manual. Instructions for use of products C8640, C8820, and C9320. *Promega Coporation*, Part# TM254.

CESG is supported by the National Institute of General Medical Sciences through the Protein Structure Initiative
NIH/NIGMS grant numbers U54 GM074901 (JLM; 07/01/05-06/30/10) and P50 GM064598 (JLM; 01/01/02-08/31/05).