CLONE AND EXPRESS PROTEIN-CODING REGIONS USING THE FLEXI® VECTOR SYSTEMS

MICHAEL SLATER, PH.D., JIM HARNETT, M.S., NATALIE BETZ, PH.D., JAMI ENGLISH, M.S., ETHAN STRAUSS, PH.D., BECKY PFERDEHIRT, B.S., AND ELAINE SCHENBORN, PH.D., PROMEGA CORPORATION

The Flexi® Vector Systems provide a rapid, efficient and high-fidelity method to transfer protein-coding regions between Flexi® Vectors with various expression or peptide-tag options while maintaining insert orientation and reading frame. A lethal gene in the Flexi® Vector cloning region allows positive selection of clones containing insert. Antibiotic resistance genes simplify selection of the desired clone during subsequent transfers to other Flexi® Vectors.

Introduction

The Flexi® Vector Systems[a,b] use two restriction enzymes, *Sgf* I[c] and *Pme* I[d], that cut infrequently in a simple yet powerful directional cloning method for protein-coding sequences. Inserts are efficiently transferred to other Flexi® Vectors, maintaining insert orientation and reading frame and eliminating the need to resequence the insert after each transfer (Figure 1). This approach is easily adapted to high-throughput formats.

Unlike site-specific recombination vector systems, the Flexi® Vector Systems do not require appending multiple amino acids to the protein of interest. The system does not require an archival entry vector and, for most applications, allows direct entry into the vector best suited to the experimental design (e.g., mammalian expression or N-terminal GST-fusion vectors).

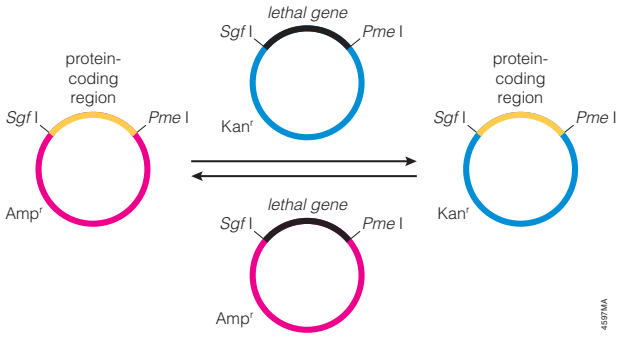


Figure 1. Transferring protein-coding regions in the Flexi® Vector Systems. The Flexi® Vector Systems employ a flexible, directional cloning method for expressing protein-coding regions with or without peptide fusion tags. The features necessary for expression and any protein fusion tag are carried on the vector backbone, and the protein-coding region can be shuttled between vectors using two rare-cutting restriction endonucleases, *Sgf* I and *Pme* I. The Flexi® Vectors contain a lethal gene, barnase, for positive selection of the protein-coding sequence and an antibiotic resistance marker for selection of colonies containing the Flexi® Vector.

The Flexi® Vector cloning strategy maintains insert orientation and reading frame, eliminating the need to resequence the insert after each transfer.

Any Flexi® Vector can act as an acceptor of a protein-coding region flanked by *Sgf* I and *Pme* I sites. All Flexi® Vectors carry the lethal barnase gene, which when replaced by the DNA fragment of interest, allows positive selection for the insert in bacterial strains commonly used for plasmid propagation and protein expression. Antibiotic resistance genes carried on the Flexi® Vectors facilitate transfer of protein-coding regions between vectors. The Flexi® Vectors contain various expression or peptide tag options that enable expression of native or fusion proteins for studying protein structure and function or protein-protein interactions (Figure 2).

ORF Capture by Flexi® Vectors Using PCR

The Flexi® Vector Systems use two infrequently cutting restriction endonucleases: *Sgf* I and *Pme* I. *Sgf* I has the fewest restriction sites in human open reading frame (ORF) sequences, and *Pme* I has the second fewest (1). These enzymes also cut infrequently in the open reading frames of many other organisms (2). Most (99%) of the annotated human open reading frames are not affected by the use of these restriction enzymes for directional cloning. However, we recommend scanning your protein-coding region for *Sgf* I and *Pme* I sites. If your protein-coding region contains these sites, consider cloning a portion of the protein-coding region or using RecA protein to protect *Sgf* I or *Pme* I sites within the protein-coding region from digestion (3). Alternatively, PCR-based, site-directed mutagenesis methods (4,5) can be used to mutate restriction enzyme sites without changing the amino acid sequence.

Sgf I and *Pme* I sites are appended to the protein-coding region by PCR. To facilitate cloning, the 5´ (forward) primer used to amplify the protein-coding region appends an *Sgf* I site, and the 3´ (reverse) primer appends a *Pme* I site (Figure 3). The *Sgf* I site is placed one base upstream of the start codon. This allows de novo initiation at the native translation start site in Flexi® Vectors that do not produce fusion proteins and allows readthrough of the *Sgf* I site in Flexi® Vectors that produce N-terminal fusion proteins. The *Pme* I site is placed at the carboxy terminus, appending a single valine residue to the last amino acid of the protein-coding

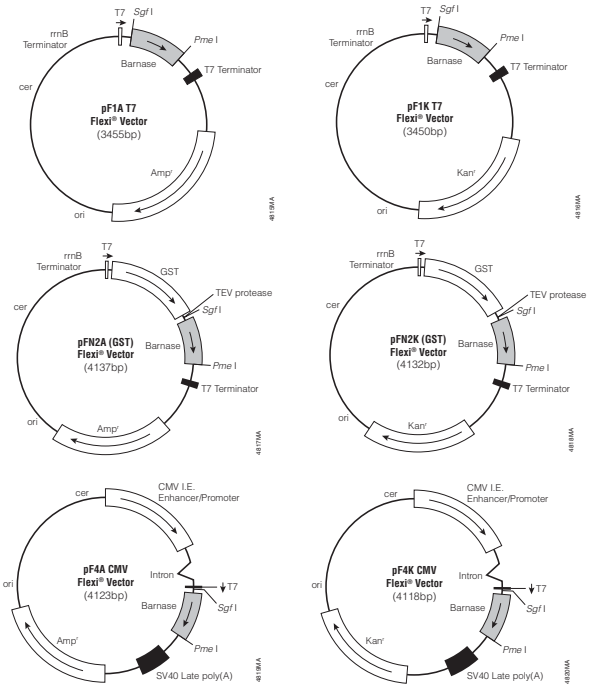


Figure 2. The Flexi® Vectors. The pF1A and pF1K T7 Flexi® Vectors[a,b] are designed for bacterial or in vitro protein expression via the T7 RNA polymerase promoter. The pFN2A and pFN2K (GST) Flexi® Vectors[a,b,e] append an N-terminal, glutathione-S-transferase (GST) peptide tag to a protein. The GST peptide tag can be removed by cleavage with TEV protease. The pF4A and pF4K CMV Flexi® Vectors[a,b,f] allow constitutive protein expression in mammalian cells using the human cytomegalovirus (CMV) intermediate-early enhancer/promoter. All of the Flexi® Vectors also contain the T7 RNA polymerase promoter for in vitro protein expression and a cer site, a 200bp site for the *E. coli* XerCD recombinase, which resolves plasmid multimers into monomers.

region. The valine codon, GTT, is immediately followed by an ochre stop codon, TAA. When a protein-coding region, flanked by *Sgf* I and *Pme* I sites, is cloned into a vector cut with *Sgf* I and *Pme* I, the translation stop codon is recreated.

Because this is a flexible system, you can design your primers to place the *Pme* I site downstream of the native stop codon, so translation terminates at the native stop codon, and a valine residue is not appended to the protein. However, by doing so you lose the ability to express the protein with a carboxy-terminal fusion tag in the future. A carboxy-terminal fusion protein can be created by fusing the blunt *Pme* I end of the protein-coding region with a blunt end generated by a different restriction enzyme (e.g., *Eco*ICR I). Primer design guidelines are provided at: **www.promega.com/techserv/tools/flexivector/**

Cloning *Renilla* Luciferase into the pFIA T7 Flexi® Vector

The synthetic *Renilla* luciferase (hRL) protein-coding region, which is codon-optimized for mammalian expression, was

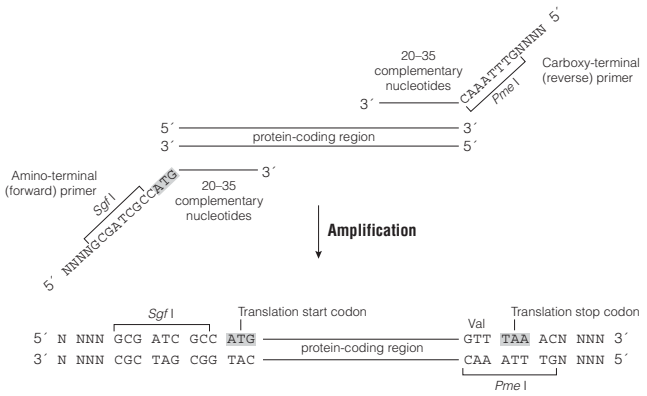


Figure 3. Amplification to append *Sgf* I and *Pme* I sites. The *Sgf* I site is upstream of the start codon. This allows de novo initiation at the start site or readthrough to append an amino-terminal peptide, depending on the vector backbone. Addition of a *Pme* I site appends a single valine codon at the 3´ end of the protein-coding region and allows either termination or readthrough to append a carboxy-terminal peptide, depending on the vector backbone.

amplified by PCR[g] using *Pfu* DNA polymerase. An overview of the cloning strategy is shown in Figure 4; the detailed protocol can be found in the *Flexi® Vector Systems Technical Manual* #TM254.

For the studies described here, the 958bp PCR product was purified, digested with the Flexi® Enzyme Blend (*Sgf* I and *Pme* I[c,d]) to yield a 941bp fragment, which was then cleaned up to remove small oligonucleotides released by the restriction enzyme digestion. Postamplification and postdigestion purifications were performed using the Wizard® SV Gel and PCR Clean-Up System (Cat.# A9281). The 941bp hRL product was ligated into the pF1A T7 Flexi® Vector that had been digested with the Flexi® Enzyme Blend. The ligation mixture was transformed into competent JM109(DE3) cells, and the cells were selected on LB plates supplemented with ampicillin.

The resultant clones were screened for *Renilla* luciferase activity and the presence of *Sgf* I and *Pme* I sites. Most (93.7%, 267/285) of the ampicillin-resistant colonies contained the hRL protein-coding region, and of those hRL positive clones, 94.5% (86/91) were cut with both *Sgf* I and *Pme* I as determined by agarose gel electrophoresis. We determined that the desired clone should be identified by screening only 3 or 4 colonies (2). Three pF1A-hRL clones were sequenced simply to confirm the insert sequence, and no unwanted variations were noted. One clone was chosen for further studies.

High-Efficiency Transfer Between Flexi® Vectors

We measured the transfer frequency between various Flexi® Vectors by screening plasmid DNA for the presence of the correctly sized *Sgf* I/*Pme* I DNA fragment encoding the synthetic *Renilla* luciferase gene. Briefly, the donor hRL-Flexi®

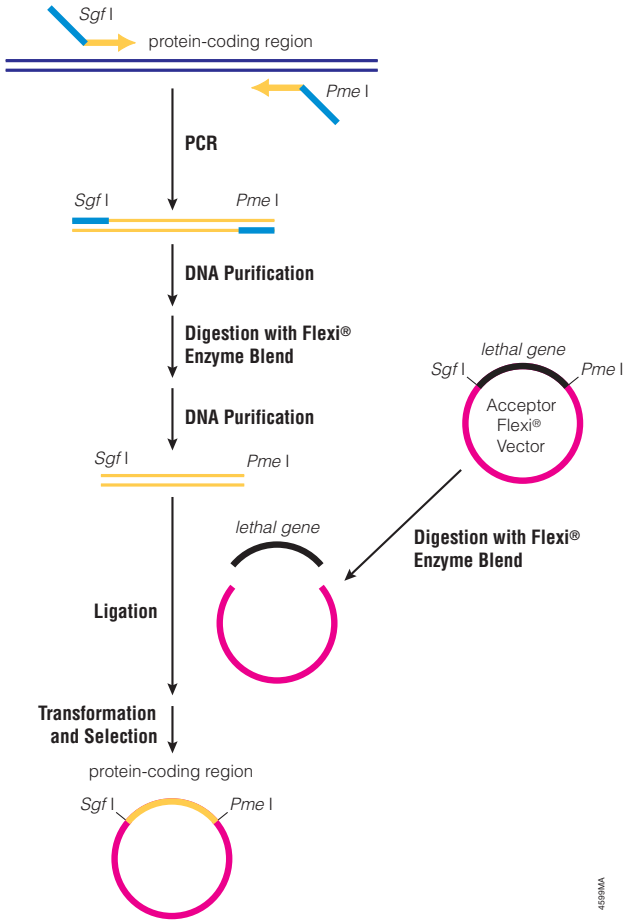


Figure 4. Cloning a protein-coding region into the Flexi® Vectors. PCR primers append *Sgf* I and *Pme* I sites onto the protein-coding region. After amplification, the PCR product is purified to remove the DNA polymerase and primers then digested with *Sgf* I and *Pme* I. The DNA is purified again to remove small oligonucleotides released by the restriction enzymes. The 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 acceptor Flexi® Vector used.

Vectors were paired with acceptor Flexi® Vectors with the opposite drug resistance. Donor and acceptor plasmid DNAs were combined, digested simultaneously with the Flexi® Enzyme Blend and ligated. Ligation reactions were transformed into JM109(DE3), and cells were selected on LB plates supplemented with the appropriate drug for the acceptor vector. The resultant clones were screened for *Renilla* luciferase activity and the presence of *Sgf* I and *Pme* I sites. Transfer frequencies between the Flexi® Vectors ranged from 88.3% to 100% with an average transfer frequency of 95% ± 4%. Again we determined that 2 to 3 clones should be adequate for screening (2).

Table 1. *Renilla* Luciferase Activity in CHO Cells Transfected with the pF4K-hRL or phRL-CMV Vector.

Vector	Average Luminescence (RLU)*
pF4K-hRL	2,806,000 ± 377,754
phRL-CMV	3,456,183 ± 768,855

*Results are expressed as the mean ± standard deviation, n = 3.

Table 2. *Renilla* Luciferase Activity in the TnT® Quick Coupled Transcription/Translation System.

Vector	Average Luminescence (RLU)
No-DNA Control	167.5 ± 12.0
pF1K-hRL	1,106,418 ± 155,552
PFN2K-hRL	292,691 ± 38,055
pF4K-hRL	838,061 ± 61,988
pRL-SV40 (Cat.# E2231)	190,797 ± 54,347
phRL-null (Cat.# E6231)	462695 ± 7,525

Delay time = 2 seconds, read time = 5 seconds. Results are expressed as the mean ± standard deviation, n = 9 for the pF1K-hRL, pFN2K-hRL and pF4K-hRL vectors, n = 3 for the pRL-SV40 and phRL-null Vectors, n = 2 for the no-DNA control.

Expression of *Renilla* Luciferase in Mammalian Cells

Chinese hamster ovary (CHO) cells were transfected with the pF4K-hRL or phRL-CMV Vector (Cat.# E6271) using the TransFast™ Transfection Reagent (Cat.# E2431). The pGL4.13[*luc2*/SV40] Vector (Cat.# E6681), which constitutively expresses firefly luciferase, was cotransfected to normalize for differences in transfection efficiency. Luciferase levels were measured using the Dual-Glo™ Luciferase Assay System (Cat.# E2920). *Renilla* luciferase levels are reported in Table 1. Firefly luciferase levels and the ratio of *Renilla* and firefly luciferase levels were similar for both pF4K-hRL and phRL-CMV vectors. Thus, protein expression levels from the pF4 Flexi® Vectors are comparable to that of other CMV-driven expression vectors.

Expression of *Renilla* Luciferase in TnT® Lysates

In vitro expression of synthetic *Renilla* luciferase for the Flexi® Vector clones was assayed with the TnT® T7 Quick Coupled Transcription/Translation System (rabbit reticulocyte lysate-based, Cat.# L1170). Plasmid DNA (500ng) was added to a 25µl TnT® reaction and incubated at 30°C for 90 minutes. Synthesized protein was labeled with the FluoroTect™ $Green_{Lys}$ in vitro Translation Labeling System (Cat.# L5001). An aliquot of the TnT® reaction (2.5µl) was added to 100µl of *Renilla* Luciferase Assay Reagent, and luminescence was measured using an Orion Microplate Luminometer (Berthold Detection Systems). Two positive controls were included: pRL-SV40 (Cat.# E2231) and phRL-null (Cat.# E6231). Results

are shown in Table 2. Proteins were separated by SDS-PAGE. The synthetic *Renilla* luciferase and *Renilla* luciferase GST-fusion proteins were of the expected sizes (data not shown). These results confirm the functionality of the T7 RNA polymerase promoter in the Flexi® Vectors for protein expression in TnT® Systems. Expression levels from different vectors were not identical; differences may be attributed to fusion protein production (e.g., pFN2K-hRL) or differing 5′ and 3′ sequences flanking the protein-coding region.

Conclusions

We have developed a new directional cloning system for protein-coding sequences based on two rare-cutting restriction enzymes, *Sgf* I and *Pme* I. The Flexi® Vector Systems provide a rapid, efficient and high-fidelity method to transfer protein-coding regions between a variety of vectors. Unlike site-specific recombination vector systems, the Flexi® Vector Systems do not require appending multiple amino acids to the amino- or carboxy-terminus of the protein or domain of interest. In addition, the systems use routine and robust molecular biology techniques and reagents.

We demonstrate efficient capture of protein-coding regions as inserts in the Flexi® Vectors following PCR. Inserts were easily transferred with >90% efficiency to other Flexi® Vectors that were digested with *Sgf* I and *Pme* I. This cloning strategy maintains insert orientation and reading frame and eliminates the need to resequence the insert after each transfer. The Flexi® Vectors enable expression of native or fusion proteins in *E. coli*, mammalian cells and in vitro translation systems. In the future, additional vectors with new expression options, such as the production of C-terminal fusion proteins, will be available. ■

References

1. National Center for Biotechnology Information (2004) RelSeq Release 6 (ftp://ftp.ncbi.nih.gov/refseq/release/)
2. Slater, M. *et al.* (2005) *Promega Notes* **89**, 11–15.
3. Schoenfeld, T., Harper, T. and Slater, M. (1995) *Promega Notes* **50**, 9–13.
4. Higuchi, R., Krummel, B. and Saiki, R.K. (1988) *Nucl. Acids Res.* **16**, 7351–67.
5. Ho, S.N. *et al.* (1989) *Gene* **77**, 51–9.

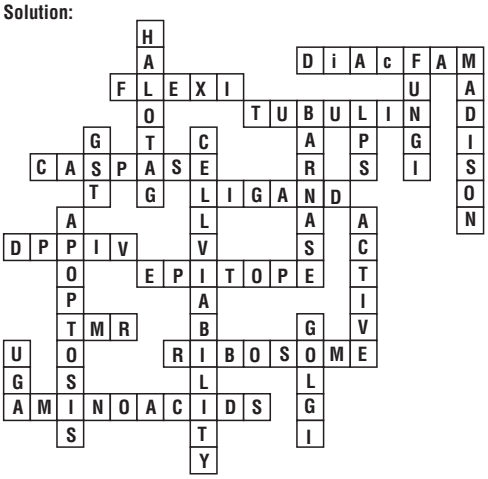
Protocol

Flexi® Vector Systems Technical Manual #TM254 (www.promega.com/tbs/tm254/tm254.html)

Just for Fun

Puzzle on page 12

Solution:



Ordering Information

Product	Size	Cat.#
pF1A T7 Flexi® Vector	20µg	C8441
pF1K T7 Flexi® Vector	20µg	C8451
pFN2A (GST) Flexi® Vector	20µg	C8461
pFN2K (GST) Flexi® Vector	20µg	C8471
pF4A CMV Flexi® Vector	20µg	C8481
pF4K CMV Flexi® Vector	20µg	C8491
Flexi® Vector System, Entry/Transfer	1 system	C8640
Flexi® Vector System, Transfer	1 system	C8820
10X Flexi® Enzyme Blend (*Sgf* I and *Pme* I)	25µl	R1851
	100µl	R1852

[a]Patent Pending.

[b]For research use only. Persons wishing to use this product or its derivatives in other fields of use, including without limitation, commercial sale, diagnostics or therapeutics, should contact Promega Corporation for licensing information.

[c]U.S. Pat. No. 5,391,487.

[d]Licensed under U.S. Pat. No. 5,945,288.

[e]This product or portions thereof is manufactured under license from Amrad Corporation Limited. For non-commercial research use only. All other uses require a license from Amrad Corporation Limited, Unit 6, 663 Victoria Street, Abbotsford, Victoria 3067, Australia, under U.S. Pat. No. 5,654,176, Australian Pat. No. 607511, Canadian Pat. No. 1338903 and other issued patents.

[f]The CMV promoter and its use are covered under U.S. Pat. Nos. 5,168,062 and 5,385,839 owned by the University of Iowa Research Foundation, Iowa City, Iowa, and licensed FOR RESEARCH USE ONLY. Commercial users must obtain a license to these patents directly from the University of Iowa Research Foundation.

[g]The PCR process is covered by patents issued and applicable in certain countries*. Promega does not encourage or support the unauthorized or unlicensed use of the PCR process.

*In the U.S., effective March 29, 2005, U.S. Pat. Nos. 4,683,195, 4,965,188 and 4,683,202 will expire. In Europe, effective March 28, 2006, European Pat. Nos. 201,184 and 200,362 will expire.

Products may be covered by pending or issued patents or may have certain limitations. Please visit our Web site for more information.