Benchmarks

level of fluorescence (Figure 2), confirming that the expression of the reporter enzyme was stable and that the selected constitutive promoter resulted in uniform expression levels. In contrast, fluorescence-activated cell sorting (FACS) analysis of an antibody expressed from the commercially available pBAD vector (Invitrogen) has shown that sub-saturating concentrations of arabinose leads to mixed populations of fully induced and fully uninduced cells (2).

The random mutation of the -10 region is a technically simple way to modulate heterologous gene expression levels for in vitro evolution experiments or other assays that require optimized expression levels, such as screens for pharmaceuticals. We chose to isolate plasmids that expressed GUS under constitutive expression conditions because such vectors could be incorporated into our high-throughput assay most conveniently. The principle virtue of this approach is that it could potentially be applied to any expression system regardless of the promoter, gene, plasmid copy number, host strain, growth media, or induction conditions. This approach will allow virtually any expression system to be adapted for a particular high-throughput assay.

REFERENCES

1. **Byrappa, S., D.K. Gavin, and K.C. Gupta.** 1995. A highly efficient procedure for site-specific mutagenesis of full-length plasmids using Vent DNA polymerase. *Genome Res.* 5:404-407.
2. **Daugherty, P.S., M.J. Olsen, B.L. Iverson, and G. Georgiou.** 1999. Development of an optimized expression system for the screening of antibody libraries displayed on the *Escherichia coli* surface. *Protein Eng.* 12:613-621.
3. **Dower, W.J., J.F. Miller, and C.W. Ragsdale.** 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* 16:6127-6145.
4. **Horwitz, M.S. and L.A. Loeb.** 1986. Promoters selected from random DNA sequences. *Proc. Natl. Acad. Sci. USA.* 83:7405-7409.
5. **Jensen, P.R. and K. Hammer.** 1998. Artificial promoters for metabolic optimization. *Biotechnol. Bioeng.* 58:191-195.
6. **Jensen, P.R. and K. Hammer.** 1998. The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl. Environ. Microbiol.* 64:82-87.
7. **Keasling, J.D.** 1999. Gene-expression tools for the metabolic engineering of bacteria. *Trends Biotechnol.* 17:452-460.
8. **King, P.V. and R.W. Blakesley.** 1986. Optimizing DNA ligations for transformation. *Focus (Gibco-BRL).* 8:1-3.
9. **Ling, M.M. and B.H. Robinson.** 1997. Approaches to DNA mutagenesis: an overview. *Anal. Biochem.* 254:157-178.
10. **Matsumura, I., J.B. Wallingford, N.K. Surana, P.D. Vize, and A.D. Ellington.** 1999. Directed evolution of the surface chemistry of the reporter enzyme β -glucuronidase. *Nat. Biotechnol.* 17:696-701.
11. **Oliphant, A.R. and K. Struhl.** 1988. Defining the consensus sequences of *E. coli* promoter elements by random selection. *Nucleic Acids Res.* 16:7673-7683.
12. **Olsen, M., B. Iverson, and G. Georgiou.** 2000. High-throughput screening of enzyme libraries. *Curr. Opin. Biotechnol.* 11:331-337.
13. **Olsen, M.J., D. Stephens, D. Griffiths, P. Daugherty, G. Georgiou, and B.L. Iverson.** 2000. Function-based isolation of novel enzymes from a large library. *Nat. Biotechnol.* 18:1071-1074.
14. **Rosenberg, M. and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* 13:319-353.
15. **Singh, M. and C. Sharma.** 1999. Detection of plasmid-encoded *gusA* gene in GUS-positive *Escherichia coli*. *BioTechniques* 26:261-264.
16. **Sutherland, J.D.** 2000. Evolutionary optimization of enzymes. *Curr. Opin. Chem. Biol.* 4:263-269.
17. **Thomas, M.R.** 1994. Simple, effective cleanup of DNA ligation reactions prior to electro-transformation of *E. coli*. *BioTechniques* 16:988-990.
18. **Youderman, P., S. Bouvier, and M.M. Susskind.** 1982. Sequence determinants of promoter activity. *Cell* 30:843-853.

We thank Profs. Brent Iverson and George Georgiou for thoughtful advice, and for the use of their fluorescence-activated cell sorter. Address correspondence to Andrew Ellington, Institute of Cellular and Molecular Biology, ICMB A4800/MBB 3.424, 26th and Speedway, University of Texas, Austin, TX 78712, USA. e-mail: andy.ellington@mail.utexas.edu

Received 29 August 2000; accepted 1 December 2000.

**Ichiro Matsumura,
Mark J. Olsen, and
Andrew D. Ellington**
*University of Texas
Austin, TX, USA*

Vectors for the Generation of FLAG[®]- or EGFP-Tagged cDNA Constructs and EGFP-Tagged Antisense RNA Constructs

BioTechniques 30:476-482 (March 2001)

Epitope or fusion protein tagging has become a common tool for the recombinant expression of proteins in mammalian cells. The addition of an N-terminal or C-terminal tag allows surveillance of the protein and facilitates identification of transfected cells and purification of the recombinant protein. Two commonly used tags are the FLAG[®] epitope (Sigma-Aldrich, St. Louis, MO, USA) and the enhanced green fluorescent protein (EGFP). The FLAG epitope consists of eight amino acid residues (DYKDDDDK). FLAG-tagged proteins can be detected with anti-FLAG antibodies in western blots and can be purified by immunoaffinity chromatography (4). EGFP is an optimized variant of the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* that emits bright green fluorescent light upon exposure to UV or blue light (2,3,5,8). EGFP-tagged protein can be detected directly in living cells without fixation or disruption. In addition, EGFP-tagged proteins can be detected in western blots or immunoprecipitated with anti-EGFP antibodies.

Eukaryotic expression vectors for the cloning of epitope-tagged cDNAs are available but are not always suited for direct cloning because they lack appropriate restriction sites or do not allow cloning of toxic genes. Certain toxic cDNAs can only be cloned and propagated in *E. coli* if leaky expression by unspecific transcription and translation is prevented (1). The use of low copy number vectors and/or vectors with several transcription termination signals can keep the nonspecific basal expression low and usually allows cloning and propagation in *E. coli*. Here, we describe the construction of vectors that allow cloning of highly toxic cDNAs in frame with an N-terminal FLAG or EGFP tag, or as EGFP-tagged antisense constructs. The cloned

Benchmarks

tagged cDNAs or antisense constructs can then be subcloned into low copy number eukaryotic expression vectors. Alternatively, a promoter can be inserted upstream of the epitope tag to drive expression of the target gene.

The pKoz vectors (Figure 1) contain the vector backbone of the pT7-SCII vector (USB, Cleveland, OH, USA), a prokaryotic expression vector designed

for the expression of highly toxic genes in *E. coli* (1). The vector backbone contains a β -lactamase (*bla*) gene for ampicillin (Amp) resistance, a Col E 1 Origin, and three sets of the *rrnB* terminators T₁ and T₂ to suppress read-through transcription from cryptic start sites on the plasmid. A 73-bp *Xba*I/*Xba*I fragment containing a Shine-Dalgarno ribosome binding site and parts

of the multi-cloning site was excised. To create pKoz/M-Flag (Figure 1), the *Xba*I sites on the vector backbone were filled in using Klenow enzyme and ligated with a blunt-end fragment that contains a Kozak consensus sequence for ribosome binding (7), a sequence to encode an N-terminal FLAG tag, and a 1578-bp stuffer insert. The FLAG tag is followed by an ATG start codon and is in frame with the ATG of the *Nde*I site. To add an N-terminal FLAG tag to a specific protein, the stuffer insert can be excised and replaced by another open reading frame using the *Nde*I site and a 3'-flanking restriction site. The FLAG-tagged cDNA can be subcloned with the Kozak sequence into a eukaryotic expression vector using 5'- and 3'-flanking restriction sites to express a recombinant protein with an N-terminal FLAG tag. We subcloned p21-activated protein kinase (PAK) (6) between the *Nde*I and *Bam*HI sites of pKoz/M-Flag to obtain PAK cDNA with an N-terminal FLAG-tag. FLAG-tagged PAK cDNA was then subcloned as a *Pac*I/*Bam*HI fragment into the inducible retroviral expression vector pRevTRE (Clontech Laboratories, Palo Alto, CA, USA). After retroviral transduction, expression of Flag-PAK was induced with doxycycline (Figure 2A).

To create pKoz/EGFP (Figure 1), the Koz/M-Flag sequence was excised with *Pac*I/*Nde*I and replaced by a *Pac*I/*Nde*I fragment that contains the Kozak consensus sequence and the open reading frame for EGFP from the pEGFP-N3 vector (Clontech Laboratories). The EGFP cDNA includes the ATG start codon, but the stop codon was removed to extend the open reading frame and be in frame with the ATG of the *Nde*I site. To create an N-terminal EGFP fusion protein, the stuffer insert can be excised and replaced by another open reading frame using the *Nde*I site and a 3'-flanking restriction site. To express a recombinant protein with an N-terminal EGFP tag, the EGFP fusion cDNA can be subcloned with the Kozak sequence into eukaryotic expression vectors using 5'- and 3'-flanking restriction sites. We subcloned PAK (6) and the p21 G-protein Cdc42 (9) between the *Nde*I and *Bam*HI sites of pKoz/EGFP to obtain EGFP-tagged PAK and Cdc42 cDNAs. EGFP-tagged PAK and Cdc42

Figure 1. Vectors for FLAG- or EGFP-tagged cDNA or EGFP-tagged antisense RNA constructs. The pKoz vectors contain three sets of *rrnB* transcription terminators to suppress leaky expression of highly toxic genes, a *bla* gene for Amp resistance, a Col E 1 origin, a Kozak consensus sequence for ribosome binding, and a stuffer insert. Multi-cloning sites (MCS1, MCS2) with unique restriction sites are located 5' and 3' of the stuffer insert. pKoz/M-Flag contains the FLAG tag 5' of the stuffer insert in frame with the *Nde*I site. pKoz/EGFP contains the EGFP tag 5' of the stuffer insert in frame with the *Nde*I site. pKoz/EGFP-anti contains the EGFP tag in antisense orientation to the stuffer insert and the Koz/M-Flag sequence. * marks a methylation-sensitive *Xba*I site. Drawings are not to scale.

Benchmarks

cDNAs were then subcloned as *PacI/BamHI* fragments into the inducible retroviral expression vector pRevTRE. After retroviral transduction, expression of EGFP-PAK and EGFP-Cdc42 was induced with doxycycline (Figure 2, B and D).

To create pKoz/EGFP-anti (Figure 1), an *XbaI/BamHI* fragment containing the Kozak sequence and the EGFP open reading frame from pEGFP-N3 was inserted between the *BamHI* and *XbaI* sites of pKoz/M-Flag. The EGFP cDNA is oriented in antisense orientation to the stuffer insert and the Koz/M-Flag sequence. To create an EGFP-tagged antisense construct, the stuffer insert can be excised with *BamHI* and a 3'-flanking restriction site. The EGFP-antisense construct can be subcloned into a eukaryotic expression vector us-

ing 5'- and 3'-flanking restriction sites to express a fusion transcript that encodes EGFP and antisense RNA of the cloned open reading frame. The expressed EGFP can be used to indirectly monitor the presence of the antisense RNA. We subcloned PAK (6) in antisense orientation to EGFP between the *BamHI* and *PacI* sites of pKoz/EGFP-anti to obtain an EGFP-tagged PAK antisense RNA construct. EGFP-tagged PAK-anti was then subcloned as a *NotI/PacI* fragment into the inducible retroviral expression vector pRevTRE. After retroviral transduction, expression of EGFP-PAK-anti was induced with doxycycline (Figure 2, C and D). Increased EGFP levels indicated expression of PAK antisense RNA and antisense RNA reduced expression of endogenous PAK.

The pKoz vectors described here can be used to obtain FLAG- or EGFP-tagged cDNA constructs, which can be subcloned into eukaryotic expression vectors for the expression of recombinant protein with N-terminal FLAG or EGFP tags. Tagged recombinant protein can be detected by western blot with anti-FLAG or anti-EGFP antibodies or in the case of EGFP by fluorescent microscopy on a single cell level. pKoz/EGFP-anti can be used to obtain EGFP-tagged antisense constructs, which can be subcloned into eukaryotic expression vectors for the expression of EGFP-tagged antisense RNA. This allows the indirect detection of antisense RNA by western blot with an anti-EGFP antibody or by fluorescent microscopy on a single-cell level. Alternatively to subcloning of tagged cDNAs from pKoz vectors into eukaryotic expression vectors, a eukaryotic promoter can be inserted into unique restriction sites 5' of the Kozak sequence. Due to a relatively low copy number and multiple sets of *rrnB* transcription terminators, the pKoz vectors are suitable for cloning of highly toxic genes such as PAK.

REFERENCES

1. Brown, W.C. and J.L. Campbell. 1993. A new cloning vector and expression strategy for genes encoding proteins toxic to *Escherichia coli*. *Gene* 127:99-103.
2. Chalfie, M., Y. Tu, G. Euskirchen, W.W. Ward, and D.C. Prasher. 1994. Green fluorescent protein as a marker for gene expression. *Science* 263:802-805.
3. Heim, R., A.B. Cubitt, and R.Y. Tsien. 1995. Improved green fluorescence. *Nature* 373: 663-664.
4. Hopp, T.P., K.S. Prickett, V.L. Price, R.T. Libby, C.J. March, D.P. Cerretti, D.L. Urdal, and P.J. Conlon. 1988. A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio/Technology* 6:1204-1210.
5. Inouye, S. and F.I. Tsuji. 1994. Aequorea green fluorescent protein. Expression of the gene and fluorescence characteristics of the recombinant protein. *FEBS Lett.* 341:277-280.
6. Jakobi, R., C.J. Chen, P.T. Tuazon, and J.A. Traugh. 1996. Molecular cloning and sequencing of the cytosolic G protein-activated protein kinase PAK I. *J. Biol. Chem.* 271: 6206-6211.
7. Kozak, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15:8125-8148.
8. Prasher, D.C., V.K. Eckenrode, W.W. Ward,

Figure 2. Expression of FLAG- and EGFP-tagged fusion proteins and EGFP-tagged antisense RNA. cDNAs for p21-activated protein kinase (PAK) and the p21 G-protein Cdc42 were subcloned into pKoz/M-Flag, pKoz/EGFP or pKoz/EGFP-anti to generate FLAG- or EGFP-tagged cDNA constructs or EGFP-tagged antisense RNA constructs. Tagged constructs were subcloned into the inducible retroviral expression vector pRevTRE and transduced into BALB3T3 fibroblasts. (A) Expression of FLAG-tagged PAK before (-Dox) and after (+Dox) induction with doxycycline was detected in a western blot with anti-FLAG antibody in transduced cells and untransduced cells (control). (B) Expression of EGFP-tagged PAK and Cdc42 after induction with doxycycline was detected in a western blot with an anti-EGFP antibody. (C) Levels of EGFP and endogenous PAK before (-Dox) and after (+Dox) induction with doxycycline were detected in western blots with anti-EGFP and anti-PAK antibodies in cells transduced with EGFP-tagged PAK antisense RNA. (D) Expression of EGFP-tagged PAK or Cdc42 and EGFP-tagged PAK antisense RNA was detected by fluorescence microscopy.

Benchmarks

- F.G. Prendergast, and M.J. Cormier. 1992. Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* 111:229-233.
9. Shinjo, K., J.G. Koland, M.J. Hart, V. Narasimhan, D.I. Johnson, T. Evans, and R.A. Cerione. 1990. Molecular cloning of the gene for the human placental GTP-binding protein Gp (G25K): identification of this GTP-binding protein as the human homolog of the yeast cell-division-cycle protein CDC42. *Proc. Natl. Acad. Sci. USA* 87:9853-9857.

This work was supported by the American Cancer Society grant no. ACS IRG 170 and a Wisconsin Breast Cancer Showhouse grant to R.J. Address correspondence to Dr. Rolf Jakobi, Dept. of Pharmacology and Toxicology, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, USA. e-mail: rjakobi@mcw.edu

Received 7 September 2000; accepted 20 November 2000.

Rolf Jakobi, Erin Moertl, and Mark A. Koepfel
*Medical College of Wisconsin
Milwaukee, WI, USA*

Targeting Random Mutations to Regions that Are Not Flanked by Existing Restriction Sites

BioTechniques 30:482-486 (March 2001)

We have developed a protocol for introducing random mutations generated by error-prone PCR into predefined regions that are not flanked by restriction sites. The method does not require any alterations outside of the mutagenized region and has a high cloning efficiency. The latter property is important for random mutagenesis because it preserves the complexity of the mutagenized PCR product. However, the method can in principle be used to clone any blunt-ended fragment. The first step of the procedure is to generate a list of blunt end-generating enzymes that do not cut the plasmid carrying the target region. This can be done using suitable sequence analysis software. The 5' and 3' flanking regions are then searched for the 5' and 3' halves, respectively, of the recognition sequences for these restriction endonucleases. When two matching halves of a recognition site are found, a restriction site is created by deleting the intervening sequence by site-directed mutagenesis. For example, if the triplets CAG and CTG occur in the 5' and 3' flanking sequences, respectively, a recognition site for *PvuII* (CAG↓CTG) can be generated by deleting the sequence between the triplets (Figure 1). The resulting unique restriction site, which makes it possible to open the plasmid exactly at the site of the deletion, can then be used for blunt-end cloning of the PCR product (3,4,6). The insert is generated by error-prone PCR (2) using the target region in the original plasmid as template and primers designed to yield a PCR product corresponding exactly to the deleted region. Before ligation, extending nucleotides at the 3' ends are removed by treating with *Pfu* DNA polymerase (5). The insertion of the PCR fragment restores the flanking regions to their original sequence.

The probability of finding the short sequences needed to use this method depends on two factors. One is how

much flexibility there is in adjusting the size of the PCR product. Since the maximum primer length is only limited by practical considerations such as the feasibility of synthesizing the oligonucleotides, the investigator has considerable freedom to adjust the length of the PCR product (minimum primer length is approximately 15 nucleotides). In most cases, there will also be some latitude in specifying the extent of the target region, which affords the investigator additional flexibility. Assuming that primers between 15 and 60 nucleotides long can be used and that the extent of the target region can be varied by 15 bp at each end, the PCR product could be adjusted by 60 bp on either side.

The other factor that determines how likely it is that this method can be used is the size of the plasmid. This affects the probability because the number of potentially useful restriction endonucleases (i.e., those without recognition sites in the original plasmid) decreases with increasing plasmid size. Enzymes with 6-bp recognition sequences are most likely to be useful because a significant number are commercially available [24 such enzymes are listed in REBASE001, not counting isoschizomers (7)]. The likelihood that a given plasmid will contain a recognition site for one of these enzymes is also less than that for enzymes with shorter recognition sequences. In addition, the two triplets needed to form a 6-bp site are easier to find than the two tetramers required for an 8-bp recognition site. Plasmids in the 5–6 kb range will lack recognition sites for approximately five to six blunt end-generating enzymes with 6-bp recognition sequences. Assuming that there are six potentially usable enzymes (with 6-bp recognition sequences) and that the insert length can be varied by 60 bp at each end, the probability of finding the sequences necessary for using this method is over 90% (calculations not shown). It is also possible to increase the number of potentially usable enzymes by removing recognition sites for enzymes that have a single site in the vector portion of the plasmid. Although this involves some additional work, it greatly improves the likelihood that this method can be used with large plasmids.

We have used this method to intro-