Simple Procedure for Creation of In-Frame Deletion Mutations Throughout an Open Reading Frame


ABSTRACT

A general method is presented for randomly mutagenizing open reading frames (ORF) to generate in-frame deletions and insertions. The protocol requires expression of the ORF of interest as a hybrid ORF-β-galactosidase fusion protein. This allows colorimetric screening for β-galactosidase activity during subsequent mutagenesis steps. Consequently, proteins with no suitable phenotypic selection or screening properties can be readily screened for mutations that disrupt and subsequently restore the reading frame of the hybrid protein. In addition, this system provides gene expression for subsequent biochemical analysis of the mutant proteins. The bovine papillomavirus type I (BPV-1) E1 ORF has been mutagenized using this method as an example.

INTRODUCTION

To determine structure/function relationships of proteins requires the ability to create discrete mutations throughout the entire coding region. Site-directed mutagenesis techniques can be used, but these approaches are generally more suitable for analysis of limited subregions (3,6,8). Mutagenesis of an entire coding region is typically performed by more generalized procedures. Commonly used techniques for generalized mutagenesis of an open reading frame (ORF) include low-fidelity polymerase chain reaction (PCR) (5) or chemical treatment (reviewed in Reference 8). Both methods create point mutations that are mapped by DNA sequencing. However, obtaining single point mutations requires conditions that generate a high percentage of wild-type sequence. Consequently, phenotypic selection or screening of the protein product is necessary to identify potential mutants among the large population of wild-type clones. For proteins with no convenient selection or screening method, this limits the applicability of such mutagenesis procedures.

To circumvent the above problems, a general mutagenesis scheme was developed that can be used to create small in-frame insertions and deletions throughout an ORF region. The method presented here is based on translational termination-linker insertion and subsequent deletion of the linker along with adjacent ORF sequences. The method
utilizes lacZ-mediated color screening to allow rapid identification of mutant clones. Mapping of mutation position is facilitated by the presence of a unique restriction endonuclease site in the inserted termination-linker sequence. This method is applicable to any ORF that can be manipulated in vitro and has been used to mutagenize the E1 ORF of bovine papillomavirus type 1 (BPV-1).

**MATERIALS AND METHODS**

**Cells and Plasmids**

*Escherichia coli* strain MC1061 [araD139 (ara,leu)7697 (lac)X74 galU galK hsr- hsm* strA] was used for all plasmid transformations and was grown on LB plates. Plasmid pGE374 was a generous gift from Dr. George Weinstock (University of Texas Health Science Center at Houston). This plasmid is pBR322-based, containing the bacterial recA promoter, translational initiation sequences and the first 35 amino acid codons of recA fused out-of-frame with lacZ. Plasmid pGE1700k was created by partial digestion of pGE1700 (10) with BamHI, followed by S1 nuclease digestion. The linker, 5'-GTACCGGTAC-3', containing a KpnI recognition sequence was ligated to gel-purified linear DNA, and blue colonies were picked and screened for the presence of a KpnI site at the E1-lacZ junction.

**Enzymes and Nucleic Acids**

DNase I (Sigma Chemical, St. Louis, MO, USA) was dissolved at a concentration of 1 mg/mL in 0.01 N HCl and stored at -70°C. All other enzymes were purchased from Boehringer Mannheim (Indianapolis, IN, USA) and used according to the manufacturer’s suggestions. Non-phosphorylated *NheI* linkers (5'-CTAGCTAGC-TAG-3'; New England Biolabs, Beverly, MA, USA) were resuspended in TE (pH 8.0) at a concentration of 1 µg/µL. The linkers were annealed by heating to 80°C for 1 min, placed in a beaker of 80°C water and slow-cooled in a 4°C refrigerator. After cooling to 4°C, the annealed linkers were stored at -20°C.

**Termination Linker Insertion**

DNase I digestion of pGE1700k. DNase I was diluted 1:1000 in 4°C 80 mM Tris-HCl (pH 7.4), 2 mM MnCl₂, 0.1 mg/mL bovine serum albumin and kept on ice. Ten micrograms of pGE1700k were adjusted to a volume of 50 µL with water, mixed with 50 µL of the diluted DNase I and incubated at room temperature. Under these conditions, DNase I cleaves both DNA strands in close proximity to each other, producing blunt-ended or short 5’ and 3’ recessed ends. Every 1.5 min, a 20-µL aliquot was removed from the reaction, and 1 µL of 0.25 M EDTA (pH 8.0) was added to stop the DNase digestion. Three microliters of each aliquot were analyzed on a 0.8% agarose gel containing 0.5 µg/mL ethidium bromide. Time points yielding 20–30% linear molecules were pooled, and the volume was adjusted to 100 µL with TE. The DNA was extracted once with phenol/chloroform/isomyl alcohol (Ambion, Austin, TX, USA) and once with chloroform. The DNA was ethanol-precipitated and resuspended in TE. DNA ends were repaired with T4 DNA polymerase in a 20-µL reaction containing 50 mM dNTPs and 1 U T4 DNA polymerase for 15 min at room temperature. Klenow polymerase (2.5 U) was added, and incubation continued another 15 min at room temperature. After extraction and ethanol precipitation, the DNA was resuspended in 25 µL of 2× ligation buffer and used for linker addition.

**Linker addition.** Linker addition was done by a modification of the procedure of Lathe et al. (4). A 50-µL reaction, containing the DNase I-digested DNA, 100 µg/mL annealed, non-phosphorylated *NheI* linkers, 1× ligase buffer (66 mM Tris-HCl, 5 mM MgCl₂, 1 mM dithiothreitol [DTT], 1 mM ATP, pH 7.5) and 2 U (Weiss units) of T4 DNA ligase, was incubated overnight at 4°C. To denature the unligated strand of the linker, the reaction was heated to 80°C for 1 min and chilled on ice. The DNA was fractionated on a 0.8% agarose gel containing ethidium bromide, and the linear DNA band was excised and purified using the Prep-A-Gene® kit (Bio-Rad, Hercules, CA, USA). The DNA was eluted with 30 µL of TE, and NaCl was then added to a final concentration of 100 mM. The ligated linkers were re-annealed as described above. Two microliters of the ligation mixture were used to transform competent MC1061. Transformed cells were plated on LB plates containing ampicillin that had been spread with 40 µL 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) (20 mg/mL). White colonies were picked and their plasmids analyzed by restriction analysis.

**Exonuclease III/S1 nuclease digestion.** DNA from clones with linker insertions in the ORF were prepared by the alkaline lysis miniprep procedure (7). Miniprep DNA (0.5–1.0 µg) was digested with *NheI* followed by ethanol precipitation. The DNA pellet was resuspended in 2.5 µL of 10 mM Tris-HCl, pH 8.0, 5 mM MgCl₂, 100 mM NaCl, 1 mM 2-mercaptoethanol (1× B buffer; Boehringer Mannheim). Exonuclease III was diluted to 2 U/µL in 1× B buffer. Equal volumes of DNA and diluted exonuclease III were mixed and incubated at room temperature for 1 min. Conditions of high NaCl, a low enzyme:DNA ratio and low reaction temperature are used to limit the extent of exonuclease III digestion (1,4,9). S1 nuclease buffer (30 M NaOAc, 50 mM NaCl, 0.03 mM ZnSO₄, pH 4.5) containing 10 U S1 nuclease in 7.5 µL was added, and the sample was incubated at 30°C for 30 min. One microliter of S1 Stop buffer (0.3 M Tris base, 50 mM EDTA, pH 8.0) was added, and the sample was incubated at 70°C for 20 min. Fragment ends were repaired by addition of 0.1 U Klenow polymerase and 1 µL 0.5 mM dNTPs, followed by incubation for 15 min at room temperature. The volume was adjusted to 100 µL with TE, and the sample was extracted and ethanol-precipitated. The DNA pellet was resuspended in 19 µL of 1× ligase buffer. T4 DNA ligase (1 U) was added, and the ligation was incubated overnight at 4°C. The ligated DNA was diluted 1:5 with TE and used for transformation of competent MC1061. One-fifth of the transformation mixture was plated on LB plates containing ampicillin and X-gal. Blue colonies were picked and analyzed using the ds-DNA Cycle Sequencing System (Life Technologies, Gaithersburg, MD, USA).
Western Blots

Whole cell extracts were prepared as previously described (10,11). Extracts were separated by electrophoresis on 10% polyacrylamide gels and transferred to nitrocellulose. Membranes were probed with anti-RecA serum (generously provided by J. Leibowitiz, University of Texas Health Science Center at Houston) and visualized using the ECL™ detection system (Amer sham, Arlington Heights, IL, USA) as described by the manufacturer.

RESULTS

A general, three-step, mutagenesis procedure was developed to create in-frame deletions throughout a target gene (Figure 1). To illustrate the procedure, plasmid pGE1700k was subjected to the mutagenesis method to create mutations in the BPV-1 E1 ORF. This plasmid was constructed to express a RecA-E1-β-galactosidase trinuclear fusion protein as required for stage 1 of the protocol. Colonies expressing pGE-1700k have a blue phenotype when plated on medium containing X-gal (not shown) and produce a protein of the correct molecular weight for the trinuclear fusion (see Figure 3).

In stage 2 of the protocol, a translation termination linker containing a unique NheI site was randomly ligated into pGE1700k. The presence of the unique restriction enzyme site in the linker sequence facilitates rapid mapping and allows discrimination between insertions in the ORF and in lacZ as shown below. A typical linker insertion

Figure 1. Outline of the mutagenesis procedure. The three stages of the mutagenesis procedure are summarized, and details of each stage are discussed in the text.
experiment yielded 500–1000 colonies, of which approximately 10% were white and thus should have the linker inserted within the E1-lacZ region. One hundred percent of the white colonies had plasmids that could be linearized with NheI, confirming insertion of the linker at a single location on the vector (data not shown). Insertion sites were mapped by digestion of mutant plasmids with BamHI plus NheI. Fragments larger than 1700 bp mapped the linker insertion site within lacZ (Figure 2, lanes 6–8), while shorter fragments mapped the insertion site within the E1 ORF (Figure 2, lanes 2–5). Plasmid DNAs from 42 clones with insertions in the E1 ORF region were sequenced (data not shown). Of this group, 36 mutants had duplications or deletions of E1 sequences flanking the linker insertion site. It is likely that these alterations occurred as a result of T4 DNA polymerase treatment of staggered ends generated during the DNase I digestion step. Elimination of the end-repair step should limit the linker addition to those molecules that were naturally blunt-ended and would thus reduce or eliminate extraneous sequence alterations.

In initial screenings, all linker insertions in plasmids from white colonies mapped to lacZ or the 3’ half of the E1 ORF. The absence of termination linkers mapping in the 5’ half of the E1 ORF might be explained by translational reinitiation downstream of a 5’ linker insertion since the E1 ORF contained multiple Shine-Dalgarno-like sequences followed by translational initiation codons in-frame with lacZ. To test this hypothesis, a termination linker was inserted at the unique Smal site in the 5’ region of the E1 ORF. All colonies had a blue-color phenotype, and plasmids isolated from 16 colonies had linker insertions at the Smal site (data not shown). We conclude that termination linker insertions in the 5’ end of the E1 ORF have a blue-color phenotype, most likely due to translational reinitiation downstream of the linker, which allows expression of low levels of β-galactosidase.

Because of the translational reinitiation problem, a kinetic assay was developed to distinguish between clones expressing low (linker insertion in the 5’ E1 ORF region) vs. high (linker insertion outside the ORF-lacZ region) levels of β-galactosidase. Blue colonies from the original transformation plating were picked and inoculated onto LB plates that had a soft agar overlay containing 0.25 mg/mL X-gal. The plates were incubated at room temperature and observed every 15 min for 1.5 h. Of 200 colonies tested, 25 showed no color formation after 1.5 h. These 25 white colonies were picked and analyzed for presence and location of the termination linker. Eight had termination linker insertions in the 5’ end of the E1 ORF, 14 in the 3’ end of the E1 ORF, and 3 in lacZ (data not shown). Clearly this approach was successful for identifying termination linker insertions in the 5’ end of the E1 ORF and should be suitable for other ORFs where translation reinitiation is a problem.

The third stage of the protocol was...
creation of in-frame deletion mutations by exonuclease III/S1 nuclease treatment after NheI digestion as described in Materials and Methods. After transformation of treated DNA, plasmids with in-frame deletions that have removed the linker and adjacent ORF sequence will yield blue colonies. To date, the deletion mutagenesis has been performed on 15 different linker-insertion mutants located in the E1 ORF. One hundred percent of blue colonies had sequence alterations that were in-frame. Several types of mutations were obtained: simple deletions, deletions plus insertions and simple insertions. The type obtained depended on whether the initial termination linker insertion was accompanied by a deletion or duplication of E1 ORF sequences and on the extent of the exonuclease III digestion. In all cases, the termination codons had been completely removed.

To confirm the phenotypic properties of the mutations at the protein level, Western blotting was performed. Figure 3 shows a representative blot of the products of the parental vector (pGE374), the starting RecA-E1-β-galactosidase hybrid expression vector (pGE1700k), a termination insertion derivative of pGE1700k and 2 independent deletion derivatives. The pGE1700k plasmid expressed an anti-RecA reactive product of the correct molecular weight that was absent from the parental plasmid. No such product was observed with the termination insertion mutation, confirming that production of the full-length fusion protein had been blocked. In both deletion mutations, an immunoreactive band with the correct molecular weight for RecA-E1-β-galactosidase was present, indicating that the termination linker had been removed and the full-length product was again being expressed.

DISCUSSION

The method described is useful for randomly mutagenizing large ORFs where no convenient screening or selection for the ORF product exits. The only requirement for implementation of this approach is expression of the ORF of interest as an ORF-LacZ fusion protein. Since there are many suitable fusion vector systems available, this method is ideal for mutagenesis of eukaryotic genes that have no distinguishable phenotype in prokaryotic cells. We have applied this procedure to the E1 ORF of bovine papillomavirus and have obtained an extensive collection of termination and in-frame deletion or insertion mutants. Characterization of the E1 mutant proteins will be presented elsewhere (J. Ludes-Meyers and V. Wilson, unpublished).

During the initial screening of white colonies for linker insertion mutations, it was noted that there were no insertions in the 5' half of the E1 ORF. This appeared to be due to translation reinitiation downstream of the termination linker insertion. Since this might be a common problem with other cloned ORFs, we developed a kinetic screening approach that more stringently distinguished insertions in the fusion protein coding region from those elsewhere in the vector sequences. Using the kinetic approach, 8 additional insertion mutations were identified in the 5' portion of the E1 ORF. The success of this approach suggests it would be a feasible solution for other ORFs that demonstrated translational reinitiation problems.
Finally, this mutagenesis approach also allows convenient biochemical characterization of the constructed ORF mutations since protein expression is an inherent requirement. After construction and analysis of the mutations, the mutant proteins can easily be freed from β-galactosidase by specific insertion of a termination linker at the ORF-lacZ junction. This results in expression of just the ORF product, which can be assayed in cell extracts or purified, if necessary, for structural and functional evaluation.

REFERENCES


This work was supported by grants from the Texas Advanced Research Program and by PHS Grant CA56699 from the National Institutes of Health. Address correspondence to Van G. Wilson, Department of Medical Microbiology and Immunology, Texas A&M University Health Science Center, College Station, TX 77843-1114, USA. Internet: v-wilson@tamu.edu

Received 14 June 1995; accepted 14 September 1995.

John H. Ludes-Meyers and Van G. Wilson
Texas A&M University Health Science Center
College Station, TX, USA