

EPICENTRE Forum

Tools & Techniques for Genomics, Proteomics & RNA Research

Identification of Permissive and Non-Permissive Peptide Insertion Sites in β -galactosidase using the EZ::TN™ In-Frame Linker Insertion Kit

Jeannie Bailey and Colin Manoil, University of Washington, Seattle, WA

Introduction

The analysis of in-frame insertions has proved to be an effective method for dissecting structure-function relationships in proteins, including the identification of permissive sites for protein engineering applications. In this report, we describe the identification of several permissive and semi-permissive sites in the *E. coli* β -galactosidase (*lacZ* gene) using the EZ::TN™ In-Frame Linker Insertion Kit. (Editor's note: An overview of this kit is presented on p. 2).

Methods

In-frame 19-codon insertions were generated in the *lacZ* gene cloned in pMLB1101, a pBR322 derivative that also encodes an ampicillin-resistance marker (*bla* gene), as described in the EZ::TN In-Frame Linker Insertion Kit instructions.

First, 0.02 picomoles of the EZ::TN™ <Not I/KAN-3> Transposon and 0.02 picomoles of the pMLB1101 plasmid (0.1 microgram) were incubated in a 10- μ l reaction mixture containing EZ::TN™ Transposase and reaction buffer for 2 hours at 37°C. EZ::TN Stop Solution was added to terminate the reaction and 1 μ l of the reaction mixture was used for electroporation of a Δ *lac recA*⁻ electrocompetent *E. coli* strain. After 30 minutes of recovery, cells were plated on rich medium containing 50 μ g/ml kanamycin to select for EZ::TN Transposon insertion clones and 40 μ g/ml XG (X-gal; 5-Bromo-4-chloro-3-indolyl β -D-galactopyranoside) and 1mM IPTG (Isopropyl β -D-thiogalactopyranoside) to screen for β -galactosidase activity. White, kanamycin-resistant colonies (denoted as "XG minus") were indicative of insertion of the 1167-bp EZ::TN <Not I/KAN-3> Transposon into

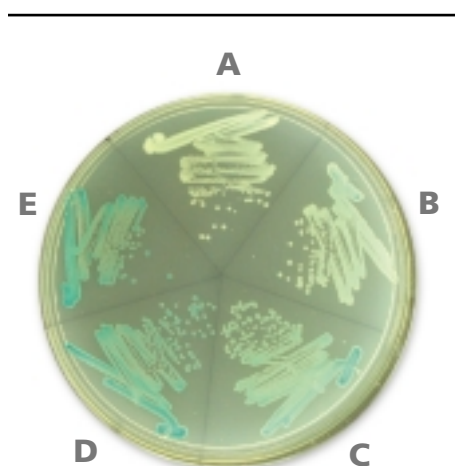


Figure 1. The growth on XG indicator agar of cells expressing different levels of β -galactosidase activity from cloned *lacZ* gene variants produced using the EZ::TN™ In-Frame Linker Insertion Kit. The plasmids carried in the different strains (clockwise from the top) are A) pBR322 (negative control), B) insertion clone D671, C) insertion clone G20, D) insertion clone L631, E) pMLB1101 (*lacZ*⁺; positive control). See Figure 4 for the exact 19-amino acid insertion site of each clone.

the *lacZ* gene of pMLB1101, with resultant "knockout" of the β -galactosidase activity. Twelve XG minus clones (putative *lacZ* insertion clones) were initially chosen for analysis. One of the 12 clones was found to be ampicillin sensitive, suggesting it had acquired an additional insertion or some other mutation in the *bla* gene of pMLB1101. Restriction digest "fingerprints" of the other 11 XG minus clones confirmed that a single EZ::TN <Not I/KAN-3> Transposon had been inserted into the *lacZ* gene. These 11 clones were used for generating 19-codon (19-amino acid) in-frame insertions by *Not* I digestion and re-ligation. Thus, each clone was independently digested with *Not* I

to release the kanamycin marker from the inserted EZ::TN <Not I/KAN-3> Transposon, re-ligated, and transformed into Δ *lac E. coli*. Colonies were selected and analyzed for growth on ampicillin agar plates containing XG and IPTG (Figure 1).

... continued on page 2

DNA Ligation in 5 minutes...
See p. 13.

In This Issue

- NEW!** 1 Identification of Permissive and Non-Permissive Peptide Insertion Sites in β -galactosidase using the EZ::TN™ In-Frame Linker Insertion Kit
- 3 Finish your DNA Sequencing Project with SequiTherm EXCEL™ II DNA Sequencing Kits
- NEW!** 4 Rapid Extraction of PCR-Ready DNA from Buccal Cells Using the BuccalAmp™ DNA Extraction Kit
- 5 High Fidelity PCR Amplification of DNA up to 40 Kb Using the MasterAmp™ Extra-Long PCR Kit
- 6 Purification of DNA from Food Materials for Detection of Genetically Modified Organisms (GMOs) Using the MasterPure™ DNA Purification Kit
- NEW!** 7 RiboShredder™ RNase Blend Destroys Unwanted RNA Quickly and Efficiently
- NEW!** 8 pIndigoBAC-5 Cloning-Ready Vectors are Extensively Tested for High Cloning Efficiency & Low Background
- 9 Three Ways to Increase Protein Yield for High Throughput Screening, Preparative Purification, or Bioprocessing
- 10 Simultaneous SNP Identification in p53 and HPV Viral Typing with Multiplex Base Excision Sequence Scanning (BESS)
- NEW!** 12 Efficiently Clone a PCR Product into the Vector of Your Choice Using the End-It™ DNA End-Repair Kit
- NEW!** 13 DNA Ligation in 5 Minutes, then Screen Your Clones in 1 Hour Using the Colony Fast-Screen™ Kit
- NEW!** 14 Produce Unbiased Libraries of Single Copy Cosmid-Sized Clones Using the EpiFos™ Fosmid Library Production Kit
- 15 Whatever the Template, You Will Never Fail at PCR Again. We Promise.

Results

Four of the 11 in-frame insertion clones (36%) analyzed for growth were medium to dark blue (XG⁺) after overnight incubation on medium containing XG and IPTG, indicative of significant levels of β -galactosidase activity. Thus, the 19-amino acids inserted into these four sites of the protein appear to be permissive based on indicator agar behavior. Assays of cellular β -galactosidase showed that these insertion derivatives expressed enzyme activity at 0.3 to 21% of the wild-type level. Two additional colonies turned medium blue after growth for about 36 hours. Two others turned only pale blue, and the final three remained white. The latter are indicative of colonies with 19-amino acid insertions into non-permissive sites of β -galactosidase.

The transposon insertion sites of DNA encoding eight independent 19-amino acid in-frame insertions were sequenced using primers provided in the EZ::TN In-Frame Linker Insertion Kit to determine the exact location of each insertion (Figure 4). The insertion sites of the four permissive clones were identified at bp 64 (G20), bp 769 (R255), bp 1898 (L631), and bp 2112 (P703). Unexpectedly, the transposon inserted in the same orientation in all four strains (with the kanamycin resistance marker and *lacZ* in the same orientation), although all three reading frames were represented. The insert at G20 is expected to be active because it falls in a region that can be deleted in β -galactosidase fusions without loss of activity. Based on the β -galactosidase three-dimensional structure¹ the L631 insertion falls at the junction between domains 3 and 4 of

the protein, which might also be expected to be relatively tolerant of insertions. The activities of the other two insertions may be explained by their evident surface exposure, although further analysis is needed.

Discussion

Our preliminary analysis of β -galactosidase insertional mutants demonstrate that the EZ::TN In-Frame Linker Insertion Kit provides a fast and efficient *in vitro* method for generating random in-frame 19-amino acid insertions into proteins encoded by cloned DNA. The results from further analyses of permissive and non-permissive sites in β -galactosidase is continuing and will be reported elsewhere.

References

1. Jacobson *et al.*, (1994) *Nature* 369:761.

An Overview of the EZ::TN™ In-Frame Linker Insertion Kit

The EZ::TN™ In-Frame Linker Insertion Kit is an *in vitro*, transposon-based protein modification system that was designed for rapid and easy production of random 19-amino acid (19-codon; 57-nucleotide) insertions into any protein encoded by a cloned DNA. Since the inserted 19 codons are read in all three reading frames, the kit can be used to identify key regulatory, binding and catalytic domains in proteins or to find permissive and non-permissive sites in proteins for protein engineering or other uses.

The overall process for introducing random 19-codon insertions into cloned DNA is shown in Figure 2. The EZ::TN <Not I/KAN-3> Transposon, containing a kanamycin-resistance marker flanked by *Not* I sites, is randomly inserted into different molecules of a target DNA, such as a clone of a protein gene, during a brief 2-hour *in vitro* reaction in the presence of EZ::TN Transposase. Following transformation of *E. coli*, a library of independent EZ::TN Transposon insertion clones is obtained. Kanamycin-resistant insertion clones for further analysis can be identified by gene functional analysis or by mapping or DNA sequencing of the transposon insertion site. Once clones are chosen, the kanamycin resistance gene is excised from the transposon by *Not* I digestion (Figure 3). Each *Not* I-digested clone is then re-ligated and re-transformed into *E. coli*. The resulting clones each contain a random 19-codon insertion that can be read in all three reading frames. Thus, the protein retains its original amino acid sequence on both sides of the insertion site.

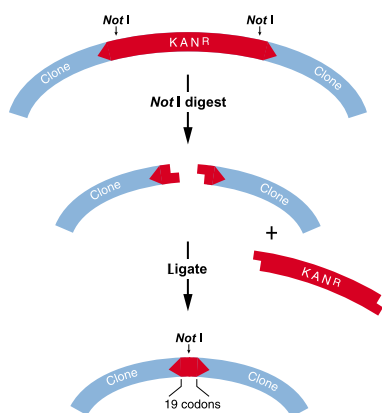


Figure 3. The EZ::TN™ <Not I/KAN-3> Transposon contains a kanamycin resistance gene flanked by *Not* I sites. A 19-codon insertion that can be read in all three reading frames is generated following *Not* I digestion and re-ligation.

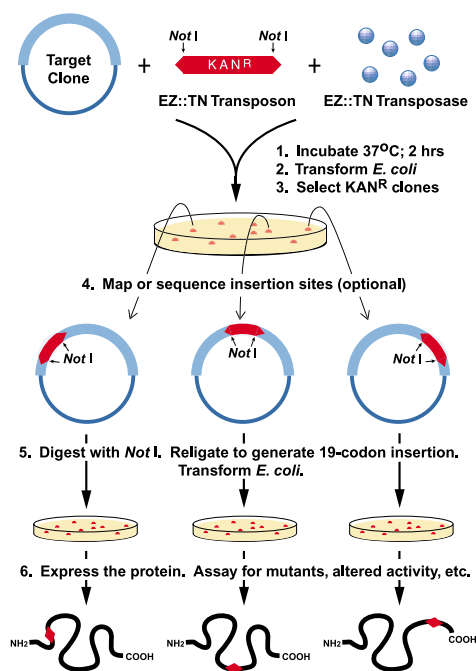


Figure 2. The EZ::TN™ In-Frame Linker Insertion Kit is based on the highly random Tn5 transposition system. A single reaction generates thousands of independent insertion clones.

Each *Not* I-digested clone is then re-ligated and re-transformed into *E. coli*. The resulting clones each contain a random 19-codon insertion that can be read in all three reading frames. Thus, the protein retains its original amino acid sequence on both sides of the insertion site.

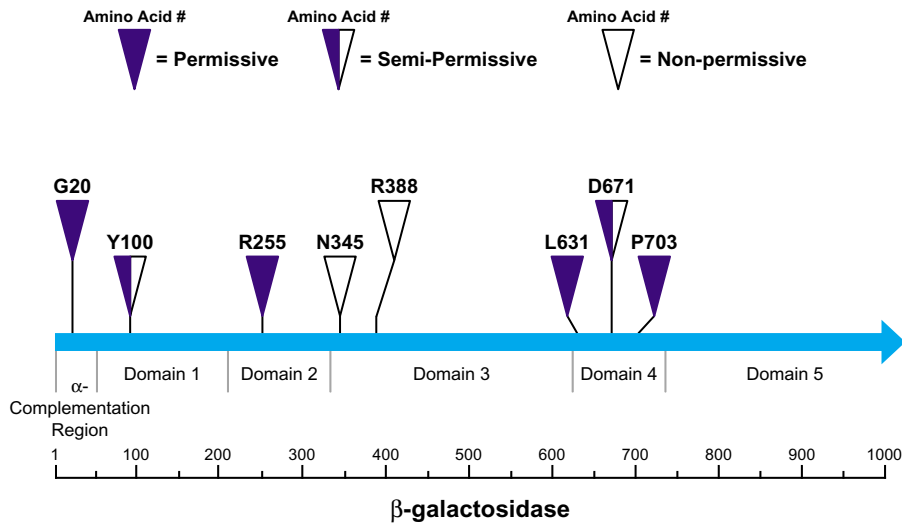


Figure 4. The location of the 19-amino acid insertion into eight independent β -galactosidase insertion derivatives. The designated amino acid is immediately N-terminal to the point of insertion. The domain structure of β -galactosidase is adapted from Jacobson *et al.*, (1994) *Nature* 369:761.

EZ::TN™ In-Frame Linker Insertion Kit
 EZI04KN 10 Reactions

TransforMax™ EC100™ Electrocompetent E. coli
 Highest efficiency electrocompetent cells available. See center insert for more information.

EC10005	5 X 100 μ l (10 electroporations)
EC10010	10 X 100 μ l (20 electroporations)

Finish your DNA Sequencing Project

Complete Difficult Regions with *SequiTherm EXCEL™ II DNA Sequencing Kits*

Sequence through hairpin loops, regions of high GC or AT content, areas of inter-strand reannealing, and inverted or direct repeats. Using end-labeled primers, *SequiTherm EXCEL II DNA Sequencing Kits* provide complete, clear sequencing data when all other methods fail.

*"I know of no other sequencing method that resolves GC compressions better than **SequiTherm EXCEL™ II.**"*

- Glenn Dawes, Affymax Research Institute, *The Scientist*, Vol.12, #6, 1998, pg. 13

SequiTherm EXCEL II resolved a "2-year-old puzzle" due to "an unreadable hairpin loop" for the Washington University genome center.

- Science, Vol. 280, 8 May 1998, pg. 816

Benefits

- Complete work on most any difficult gap to finish your sequencing project.
- High flexibility - you can perform both cycle and isothermal sequencing with one kit.
- Kits compatible with LI-COR® and ALF™ DNA Sequencers (both dye-primer and internal label) and with ABI™ automated sequencers (dye-primer).
- Reduced background signal.

G A T C

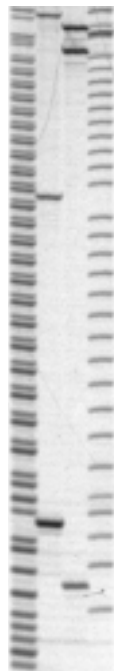


Figure. Sequencing through a GC-rich trinucleotide repeat. Supercoiled plasmid template containing (CGG)₂₃ was sequenced using the *SequiTherm EXCEL II* isothermal sequencing protocol.

SequiTherm EXCEL™ II DNA Sequencing Kit - LC For 25-41 cm gels on LI-COR® Sequencers

SE9101LC (100 Sequences)

SequiTherm EXCEL™ II DNA Sequencing Kit - LC For 66 cm gels on LI-COR® Sequencers

SE9202LC (100 Sequences)

SequiTherm EXCEL™ II Long-Read™ DNA Sequencing Kit - ALF™

For ALF DNA Sequencers
 SE8301A (100 Sequences)

SequiTherm EXCEL™ II DNA Sequencing Kit

(for manual cycle and non-cycle sequencing)

SEM79020	20 Sequences
SEM79050	50 Sequences
SEM79100	100 Sequences