

## 25

---

# Mutagenic PCR of Protein-Coding Genes for In Vitro Evolution

Ichiro Matsumura and Andrew D. Ellington

## 1. Introduction

In vitro protein evolution is an efficient approach for study of the structure and function of a protein, or to enhance its industrial utility (*I*). One round of evolution consists of random mutation of a protein-coding gene, expressing the resulting library in a population of micro-organisms, and high-throughput screening or selection of clones that most strongly exhibit a desired phenotype (“winners”). After many rounds, mutations that confer the phenotype accumulate on a single allele, e.g., the authors have isolated an octuple mutant of the *Escherichia coli*  $\beta$ -glucuronidase with catalytic activity resistant to roughly 80-fold higher concentrations of glutaraldehyde than that of the wild-type enzyme (*2*). Here we describe a variation of the mutagenic polymerase chain reaction (PCR) (*3,4*), and explanation of why the authors prefer this method to other random mutagenesis techniques.

The expression system and high-throughput assay should be developed prior to any random mutagenesis. The latter has utility only in the context of a screen that is precise, high-throughput, and sensitive enough to detect the desired activity in the ancestral protein, under stringent assay conditions. The following expression/assay system is used as an example: *E. coli* colonies, transformed with the wild-type  $\beta$ -glucuronidase expression vector, *gusA*-pBS $\Delta$ , turn green when induced on Luria-Bertani (LB) plates containing 5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucuronide (X-gluc). The stringency of this assay can be increased by pretreating the colonies with glutaraldehyde before incubation with X-gluc (*2*).

From: *Methods in Molecular Biology*, vol. 182: *In Vitro Mutagenesis*, 2nd ed.  
Edited by: J. Braman © Humana Press Inc., Totowa, NJ

In mutagenic PCR, manganese (Mn) ions, excess magnesium (Mg), and a nucleotide imbalance (4), or synthetic nucleotide analogs (5), are added to a PCR reaction to decrease the fidelity of Taq polymerase. Sequence diversity can also be generated by mutator strains (6), hypermutagenesis (7), cassette mutagenesis (8), DNA shuffling (9), or family shuffling (10). The virtues and disadvantages of each method is summarized in **Table 1**. The authors prefer mutagenic PCR (4) when mutating whole genes, because it allows a series of libraries to be made, which differ in mutation frequency simply by altering PCR conditions, which is important, because libraries with too many wild-type or overmutated clones decrease the effective throughput of the screen. The appropriate mutation frequency is theoretically dependant on the length of the gene, the average number of random mutations that the protein can accept without unfolding, the mutation bias, and the throughput of the screen. This chapter shows that, in practice, the simplest way to evaluate a library of random mutants is to utilize the high-throughput assay under permissive conditions.

One drawback of mutagenic PCR is that not all types of nucleotide changes occur at equal frequency (11). The authors sometimes (2), although not always (12), observe a significant transition bias among winners. All PCR-generated libraries have similar transition biases, and selection will magnify or diminish this bias, depending on the nature of the beneficial mutations. Mutagenic PCR, using the commercially available nucleotide analogs 8-oxo-deoxyguanosine triphosphate (dGTP) and dPTP, rather than Mn, can alter the bias and further increase sequence diversity (5). Cassette mutagenesis, based on oligonucleotides with random codons, does not exert a codon bias, but only permits randomization of a small number of nearby amino acid residues (8). Miyazaki and Arnold (13) have used cassette mutagenesis to randomize codons that were identified by sequencing winners generated by mutagenic PCR.

The mutation rates of the other methods are more difficult to control. In vitro random recombination by DNA shuffling is attractive, because it can theoretically unite all beneficial mutations from a screen in a single round. In practice, however, the 0.7% mutation rate associated with DNA shuffling (9) is too high for many applications, which necessitates multiple rounds of shuffling and screening (2,14-17), or labor-intensive procedures, to reduce the mutation rate (18). The mutation rate of the family shuffling process has not been as well-characterized, but thus far seems even higher (10).

Propagation of the vector in a mutator strain, such as XL-1 Red (6), is also attractive, because this can obviate the labor of subcloning. The latter procedure induces a maximum mutation rate of 0.0005 mutations/bp, after propagation of a high-copy-number plasmid for 30 generations of bacterial growth, which is somewhat lower than the authors' target. Mutagenic PCR is therefore appropriate for most experiments in which the relationship between protein

Table 1

Au: Pls  
define dPTP

**Table 1**  
**Comparison of Random Mutagenesis Methods**

| Method                            | Principle   | Rate (%) | Bias                            | Application                                       | Disadvantages                                     | Ref. |
|-----------------------------------|---|----------|---------------------------------|---|---|------|
| Mutator strain XL-1 RED           | <i>E. coli</i> lacking 3 DNA repair functions           | 0.05     | None, according to manufacturer | Easy to obtain large libraries                    | Mutation rate low                                 | (6)  |
| Mutagenic PCR (Mn <sup>2+</sup> ) | Low-fidelity PCR  | 0.02–0.7 | Transition ( <i>see text</i> )  | Mutation frequency can be adjusted                | <i>See text.</i>                                  | (4)  |
| Mutagenic PCR (dPTP, 8-oxo-dGTP)  | Low-fidelity PCR  | 0.02–0.7 | Transition ( <i>see text</i> )  | Different mutation bias than Mn <sup>2+</sup> PCR | Requires nucleotide analogs                       | (5)  |
| Hypermutagenesis                  | Low-fidelity reverse transcription                      | 0.5–10   | Transition                      | High mutation rate for up to 200 bp               | Mutation rate too high for whole-gene mutagenesis | (7)  |
| DNA shuffling                     | Random fragmentation and PCR-like reassembly of winners | 0.7      | Transition                      | Brings beneficial mutations together              | High mutation rate difficult to control           | (9)  |
| Family shuffling                  | DNA shuffling of natural homologs                       | >0.7?    | Transition?                     | Randomly recombines divergent homologs            | High mutation rate, technically difficult         | (10) |
| Cassette mutagenesis              | Extension of oligonucleotides with random nucleotides   | 0–75     | None                            | Randomizes 1–8 nearby codons                      | Requires structural model                         | (8)  |

structure and function is unclear. The following protocol details the random mutation of a gene, the subcloning and transformation of the resulting library, and the evaluation of the library using a high-throughput assay.

## 2. Materials

### 2.1. Reagents and Buffers

1. DNA oligonucleotides can be purchased commercially from GenSet (La Jolla, CA) or Operon (Alameda, CA), and stored as a 100  $\mu\text{mol}$  concentrated stock solution in 10  $\text{mM}$  Tris-HCl, pH 9.5, at  $-20^\circ\text{C}$ . These can be diluted in water to 1/10  $\mu\text{mol}$  stocks, and stored at  $-20^\circ\text{C}$ . The primers should contain complementary region of primers, so that the melting temperature ( $T_m = 4^\circ\text{C} \times [G + C] + 2^\circ\text{C} \times [A + T]$ ) is  $75\text{--}80^\circ\text{C}$ . Primers should include restriction sites at the 5' ends so that the gene can be subcloned into the expression vector, to express the gene in-frame, plus four or more additional nucleotides at the extreme 5' end.
2. Buffered deoxynucleotide triphosphates (dNTPs) (Amersham Pharmacia Biotech, Piscataway, NJ), should be mixed and diluted to 4  $\text{mM}$  (each dNTP) in water, and stored at  $-20^\circ\text{C}$ .
3. 5X normal PCR buffer: 300  $\text{mM}$  Tris-HCl, pH 8.5, 75  $\text{mM}$   $(\text{NH}_4)_2\text{SO}_4$ , 10  $\text{mM}$   $\text{MgCl}_2$ . If PCR yields are low, the DNA optimizer kit, containing a series of alternative buffers, can be purchased from Invitrogen (Carlsbad, CA).
4. Agarose (for 1–10-kb DNA fragments): 0.8% LE agarose (FMC, Rockland, ME) in 1X TAE (40  $\text{mM}$  Tris-acetate, 1  $\text{mM}$  EDTA), 0.5  $\mu\text{g}/\text{mL}$  ethidium bromide.
5. 10X mutagenic addition buffer. The mutagenic PCR protocol is based on that of Cadwell and Joyce (3). Their buffer differs from normal (nonmutagenic) PCR buffer, in that it contains final (1X) concentrations an additional 0.8  $\text{mM}$  deoxythymidine triphosphate (dTTP), 0.8 deoxycytidine triphosphate (dCTP), 4.8  $\text{mM}$   $\text{MgCl}_2$ , 0.5  $\text{mM}$   $\text{MnCl}_2$ . The authors therefore make a 10X cocktail of 8  $\text{mM}$  dTTP, 8 dCTP, 48  $\text{mM}$   $\text{MgCl}_2$ , 5  $\text{mM}$   $\text{MnCl}_2$  that can be added to otherwise normal *Taq* PCR reactions to make them mutagenic. Cadwell and Joyce (4) add the  $\text{MnCl}_2$  last, to prevent precipitation, but the authors do not do this, and have not had this problem. The cocktail is stored at  $-20^\circ\text{C}$ .
6. Qiaquick PCR purification and gel extraction kits were purchased from Qiagen (Chatsworth, CA), and stored at room temperature.
7. 5X Gibco-BRL T4 DNA ligase buffer: 250  $\text{mM}$  Tris-HCl, pH 7.6, 50  $\text{mM}$   $\text{MgCl}_2$ , 25% (w/v) polyethylene glycol 8000, 5  $\text{mM}$  adenosine triphosphate (ATP), 5  $\text{mM}$  dithiothreitol (DTT) (19). This buffer comes with the enzyme, if purchased from Gibco-BRL Life Technologies (Bethesda, MD). Store at  $-20^\circ\text{C}$ .
8. SOC: Mix (per L) 20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 2.5  $\text{mM}$  KCl. Autoclave for 20 min ( $121^\circ\text{C}$ , 14 atmospheres, slow cooling). Autoclave 1  $\text{M}$   $\text{MgCl}_2$  and 20% (w/v) glucose, separately, and add 10  $\text{mM}$   $\text{MgCl}_2$  and 20  $\text{mM}$  glucose to the media, after it has cooled to room temperature.
9. TB buffer: Mix 10  $\text{mM}$  PIPES, 15  $\text{mM}$   $\text{CaCl}_2$ , 250  $\text{mM}$  KCl; pH to 6.7, using KOH, then add 55  $\text{mM}$   $\text{MnCl}_2$ , and filter-sterilize, and store at  $4^\circ\text{C}$  (20).

## Mutagenic PCR of Protein-Coding Genes

265

10. Chemically competent cells (20):
  - a. Propagate 50 mL *E. coli* cells to mid-log ( $OD_{600} \sim 0.5$ ) at 18°C (room temperature works almost as well). Incubate on ice for 10 min or more.
  - b. Harvest cells by centrifuging at 2500g (3000 rpm in Beckman J-6B centrifuge) for 10 min at 4°C.
  - c. Resuspend gently in a small volume of ice cold TB, bring up to 16 mL, and spin down again.
  - d. Resuspend in 3.72 mL TB, add 0.28 mL dimethylsulfoxide (DMSO), aliquot 20 × 0.2 mL portions into microcentrifuge tubes in powdered dry ice.
  - e. Store competent cells at -80°C.
11. Butterfly nitrocellulose membranes (Schleicher and Schuell, Keene, NH).
12. LB agar plates should be poured the evening before the transformation, and allowed to cool at room temperature overnight:
  - a. Mix (per L) 10 g tryptone, 5 g yeast extract, 10 g NaCl, 15 g agar. Autoclave for 20 min (121°C, 14 atmospheres, slow cooling).
  - b. Mix the following stocks in sterile water: 100 mg/mL ampicillin (AMP), 25 mg/mL kanamycin (KM), 100 mM isopropylthio-β-D-galactoside (IPTG). These can be stored at -20°C. Mix 50 mg/mL X-gluc in DMSO; this cannot be stored long; do not mix more than need.
  - c. Let the LB agar media cool to 50°C (when flask can be held in one's bare hands). Add 1 mL AMP and 1 mL KM stocks/L LB agar plates. For some plates, also add 0.5 mM IPTG and 0.08% X-gluc. Unused plates can be stored at 4°C, but the X-gluc in plates is unstable, so they should be used within 1 wk.

### 2.2. Enzymes

All enzymes can be obtained from the indicated manufacturer, and stored at -20°C.

1. *Taq* polymerase: Perkin-Elmer (Foster City, CA) or Promega (Madison, WI).
2. Proteinase K: Boehringer-Mannheim (Indianapolis, IN). Mix small volumes (<100 μL) of 10 mg/mL stock solution in H<sub>2</sub>O, and store at -20°C. Mix new stock, if precipitation is visible after thawing.
3. Restriction enzymes (New England Biolabs [NEB], Beverly, MA).
4. Shrimp alkaline phosphatase (Boehringer-Mannheim, Indianapolis, IN).
5. T4 DNA ligase (Gibco-BRL Life Technologies, Bethesda, MD). Just prior to use, add 1 μL of T4 DNA ligase (1 U) into 9 μL of 1X ligase buffer (1/10 dilution). Note that NEB uses a different unit definition of ligase activity, and does not provide a buffer containing polyethylene glycol.

### 2.3. Expression System

1. *gusA*-pBSΔ is an expression vector that uses the *lac* promoter to drive the expression of β-glucuronidase. It is based on pBluescript (Stratagene, La Jolla, CA), which confers AMP resistance (2).
2. The InvαF'/pREP4 strain (2) expresses little endogenous β-glucuronidase activity. The pPREP4 plasmid (Qiagen, Chatsworth, CA) is a constitutive *lac* repres-

sor expression vector that stabilizes *gusA*-pBSΔ. It is based on pACYC184, which confers resistance to 25 μg/mL KM.

### 3. Method

#### 3.1. Mutagenic PCR

1. Set up a 50 μL PCR reaction containing 1X normal PCR buffer, 200 μmol dNTPs, 100–500 nM primers, 20 fmol template, and 0, 0.312, 0.625, 1.25, 2.5, or 5 μL 10X mutagenic addition buffer (*see Note 1*). Add water to 49 μL, then 1 μL (2.5 U) *Taq* polymerase.
2. Overlay a drop of light mineral oil, and amplify for 25 cycles of PCR. For primers with 75–80°C  $T_m$ s, 25× (94°C for 30 s, 72°C for 1 min/kb of desired product).
3. Check yield by running a microliter of each reaction in an agarose gel, with a known quantity of a mol wt standard. The total yield should be 1–10 μg (*see Note 2*).
4. Eliminate *Taq* polymerase. Add EDTA to 5 mM, sodium dodecyl sulfate to 0.5%, and proteinase K to 50 μg/mL, and incubate at 65°C for 15 min (*see Note 3*).
5. Purify PCR product, using Qiagen Qiaquick PCR purification kit, as directed by the manufacturer (*see Note 4*).

#### 3.2. Subcloning

1. Set up 50-μL restriction digests: 1–10 μL of each purified PCR product, or 2 μg of the expression vector, separately, with the appropriate restriction buffer and an excess of restriction enzyme overnight (*see Note 5*).
2. Confirm digestion by running 1 μL of each reaction in an agarose gel, with a known quantity of a mol wt standard.
3. Dephosphorylate the vector by adding 1 μL of shrimp alkaline phosphatase to the restriction digest, and incubate at 37°C for another hour (*see Note 5*).
4. Gel purify the insert and vector:
  - a. Pour a 0.8% agarose gel with one large well.
  - b. Add loading dye to each digest, and run in separate gels.
  - c. Visualize bands on UV light box, and excise desired bands, using a clean razor blade.
  - d. Purify agarose-encased DNA, using the Qiagen Qiaquick gel extraction kit, as directed by the manufacturer (*see Note 4*).
5. Check the purification yield by running 1 μL of each reaction in an agarose gel, with a known quantity of a mol wt standard.
6. Ligate 20 fmol purified vector with 60 fmol insert (*see Note 6*) in 20-μL reactions with 1X BRL T4 DNA ligase buffer + 0.1 U T4 DNA ligase (*19*). Set up one ligation for each library, plus control reactions without insert, and without insert or ligase. Incubate at 16°C overnight.

#### 3.3. Determination of Inactivation Rate (Using β-Glucuronidase as Example)

1. Transform library into chemically competent *E. coli* (*20*) (*see Note 7*):
  - a. Treat kill ligase by incubating each ligation at 65°C for 10 min.

## Mutagenic PCR of Protein-Coding Genes

267

- b. Thaw frozen competent cells at room temperature, and place on ice after thawing has started.
  - c. Add equal volumes of each ligation to competent cells, so that no 200- $\mu$ L aliquot of cells receives more than 10 ng total DNA. Incubate on ice for 30 min.
  - d. Heat shock cells by incubating tubes at 42°C for 40 s. Place tubes back on ice briefly, and add 800  $\mu$ L SOC. Incubate at 37°C for 1 h.
  - e. Plate serial dilutions of each transformation onto noninducing LB agar AMP/KM plates, in order to estimate the size of each library, and to propagate each library at an appropriate colony density for the high-throughput screen.
2. Assess the enzyme activity of each clone in the high-throughput screen, under permissive conditions. Adsorb the transformed colonies to a nitrocellulose filter, and transfer them (colony side up) to a second LB agar AMP/KM/IPTG/X-gluc plate, and incubate for 2 h at 37°C. The authors generally look for libraries that contain 70–99% clones exhibiting less than wild-type activity (*see Note 8*).

---

Au: Pls  
define SOC

---

### 4. Notes

1. Each of these should yield a library mutated at a different rate, although the authors do not expect a linear relationship between Mn concentration and mutation rate (**21**).
2. If PCR in **step 2** fails, try the different PCR buffers in the PCR Optimizer kit in conjunction with 0 and 0.625  $\mu$ L 10X mutagenic addition buffer. PCRs of longer templates (> 2 kb) generally do not amplify at higher concentrations of Mn, but these should not be mutated at such high rates.
3. *Taq* polymerase interferes with cloning by filling in restriction digested DNA. It binds tightly to the ends of PCR products, and cannot be eliminated through silica chromatography alone (**22**).
4. The Wizard PCR prep kit (Promega) and the GeneClean DNA purification kit (Bio101, Vista, CA) should also work.
5. In order to reduce the background of colonies transformed with uncut vector, the authors prefer long double digestions of small amounts of vector, followed by dephosphorylation using the thermolabile shrimp alkaline phosphatase.
6. 20 fmol of a 1-kb fragment of DNA is 13 ng (**19**).
7. If the library size is smaller than the throughput of the screen, try electroporation: Denature T4 DNA ligase by heating reaction for 65°C for 10 min. Remove salt, as follows. Add 30  $\mu$ L H<sub>2</sub>O and 500  $\mu$ L butanol to each ligation. Vortex, and precipitate in microcentrifuge (13,000g) for 10 min. Carefully remove supernatant, air-dry pellet, and resuspend in 5  $\mu$ L H<sub>2</sub>O (**23**). Electroporate each ligation into *E. coli* (**24**). Plate small portions of each transformation (100  $\mu$ L, diluted 1/10, 1/100, 1/1000) onto LB agar AMP/KM plates.
8. The authors conservatively consider any colony less green than the wild-type control to be inactivated. Most random mutations, particularly transitions, will be silent or selectively neutral (**25**). The actual numbers of mutations/allele will follow a Poisson distribution, so that the variation among clones within any library will be great (**4**). The authors' experience so far is that libraries containing

50–99% inactivated clones will yield winners for screens of 1000–10,000 clones. Higher frequencies may be advantageous for higher-throughput screens and selections ( $>10^6$ ) (26).

## References

1. Sutherland, J. D. (2000) Evolutionary optimisation of enzymes. *Curr. Opin. Chem. Biol.* **4**, 263–269.
2. Matsumura, I., Wallingford, J. B., Surana, N. K., Vize, P. D., and Ellington, A. D. (1999) Directed evolution of the surface chemistry of the reporter enzyme beta-glucuronidase. *Nat. Biotechnol.* **17**, 696–701.
3. Cadwell, R. C. and Joyce, G. F. (1992) Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* **2**, 28–33.
4. Cadwell, R. C. and Joyce, G. F. (1994) Mutagenic PCR. *PCR Methods Appl.* **3**, S136–S140.
5. Zacco, M., Williams, D. M., Brown, D. M., and Gherardi, E. (1996) An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol.* **255**, 589–603.
6. Greener, A. and Callahan, M. (1994) XL1-Red: a highly efficient random mutagenesis strain. *Strategies (Stratagene)* **7**, 32–34.
7. Martinez, M. A., Pezo, V., Marliere, P., and Wain-Hobson, S. (1996) Exploring the functional robustness of an enzyme by in vitro evolution. *EMBO J.* **15**, 1203–1210.
8. Black, M. E., Newcomb, T. G., Wilson, H. M., and Loeb, L. A. (1996) Creation of drug-specific herpes simplex virus type 1 thymidine kinase mutants for gene therapy. *Proc. Natl. Acad. Sci. USA* **93**, 3525–3529.
9. Stemmer, W. P. (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **91**, 10,747–10,751.
10. Crameri, A., Raillard, S. A., Bermudez, E., and Stemmer, W. P. (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**, 288–291.
11. Fromant, M., Blanquet, S., and Plateau, P. (1995) Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal. Biochem.* **224**, 347–353.
12. Matsumura, I. and Ellington, A. D. (1999) In vitro evolution of thermostable p53 variants. *Protein Sci.* **8**, 731–740.
13. Miyazaki, K. and Arnold, F. H. (1999) Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. *J. Mol. Evol.* **49**, 716–720.
14. Crameri, A., Whitehorn, E. A., Tate, E., and Stemmer, W. P. (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol.* **14**, 315–319.
15. Crameri, A., Dawes, G., Rodriguez, Jr., E., Silver, S., and Stemmer, W. P. (1997) Molecular evolution of an arsenate detoxification pathway by DNA shuffling. *Nat. Biotechnol.* **15**, 436–438.

*Mutagenic PCR of Protein-Coding Genes*

269

16. Stemmer, W. P. (1994) Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389–391.
17. Zhang, J. H., Dawes, G., and Stemmer, W. P. (1997) Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc. Natl. Acad. Sci. USA* **94**, 4504–4509.
18. Zhao, H. and Arnold, F. H. (1997) Optimization of DNA shuffling for high fidelity recombination. *Nucleic Acids Res.* **25**, 1307–1308.
19. King, P. V. and Blakesley, R. W. (1986) Optimizing DNA ligations for transformation. *Focus (Gibco-BRL)* **8**, 1–3.
20. Inoue, H., Nojima, H., and Okayama, H. (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23–28.
21. Beckman, R. A., Mildvan, A. S., and Loeb, L. A. (1985) On the fidelity of DNA replication: manganese mutagenesis in vitro. *Biochemistry* **24**, 5810–5817.
22. Wybranietz, W. A. and Lauer, U. (1998) Distinct combination of purification methods dramatically improves cohesive-end subcloning of PCR products. *Biotechniques* **24**, 578–580.
23. Thomas, M. R. (1994) Simple, effective cleanup of DNA ligation reactions prior to electro- transformation of *E. coli*. *Biotechniques* **16**, 988–990.
24. Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* **16**, 6127–6145.
25. Moore, J. C. and Arnold, F. H. (1996) Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nat. Biotechnol.* **14**, 458–467.
26. Zaccolo, M. and Gherardi, E. (1999) The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase. *J. Mol. Biol.* **285**, 775–783.