1-1-1997

Long-inverse PCR to generate regional peptide libraries by codon mutagenesis

Dominic P. Eisinger
Dartmouth College

Bernard L. Trumpower
Dartmouth College

REFERENCES

1. **Beavis, R.C. and B.T. Chait.** 1990. Rapid, sensitive analysis of protein mixtures by mass spectrometry. Proc. Natl. Acad. Sci. USA *87*:6873-6877.
2. **Billeci, T.M. and J.T. Stults.** 1993. Tryptic mapping of recombinant proteins by matrix-assisted laser desorption/ionization mass spectrometry. Anal. Chem. *65*:1709-1716.
3. **Bornsen, K.O., M.D. Mohr and H.M. Widmar.** 1995. Ion exchange and purification of carbohydrates on a Nafion® membrane as a new sample pretreatment for matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun. Mass Spectrom. *9*:1031-1034.
4. **Hutchens, T.W. and T.-T. Yip.** 1993. New desorption strategies for the mass spectrometric analysis of macromolecules. Rapid Commun. Mass Spectrom. *7*:576-580.
5. **Mock, K.K., C.W. Sutton and J.S. Cottrell.** 1992. Sample immobilization protocols for matrix-assisted laser-desorption mass spectrometry. Rapid Commun. Mass Spectrom. *6*:233-238.
6. **Moritz, R., J. Eddes, H. Ji, G.E. Reid and R.J. Simpson.** 1995. Rapid separation of proteins and peptides using conventional silica-based supports: identification of 2-D gel proteins following in-gel proteolysis, p. 311-319. *In* J.W. Crabb (Ed.), Techniques in Protein Chemistry VI. Academic Press, San Diego.
7. **Ortiz, M.L., M. Calero, C.F. Patron, L. Castellanos and E. Mendez.** 1992. Imidazole-SDS-Zn reverse staining of proteins in gels containing or not SDS and microsequence of individual unmodified electroblotted proteins. FEBS Lett. *296*:300-304.
8. **Patterson, S.D.** 1995. Matrix-assisted laser-desorption/ionization mass spectrometric approaches for the identification of gel-separated proteins in the 5–50 pmol range. Electrophoresis *16*:1104-1114.
9. **Patterson, S.D. and V. Katta.** 1994. Prompt fragmentation of disulfide-linked peptides during matrix-assisted laser desorption ionization mass spectrometry. Anal. Chem. *66*:3727-3732.
10. **Shevchenko, A., M. Wilm, O. Vorm and M. Mann.** 1996. Mass-spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal. Chem. *68*:850-858.
11. **Vorm, O., P. Roepstorff and M. Mann.** 1994. Improved resolution and very high-sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. Anal. Chem. *66*:3281-3287.
12. **Xiang, F. and R.C. Beavis.** 1994. A method to increase contaminant tolerance in protein matrix-assisted laser desorption/ionization by the fabrication of thin protein-doped polycrystalline films. Rapid Commun. Mass Spectrom. *8*:199-204.
13. **Zaluzec, E.J., D.A. Gage, J. Allison and J.T. Watson.** 1994. Direct matrix-assisted laser-desorption ionization mass-spectrometric analysis of proteins immobilized on nylon-based membranes. J. Am. Soc. Mass Spectrom. *5*:230-237.
14. **Zhang, H., P.E. Andren and R.M. Caprioli.** 1995. Micro-preparation procedure for high-sensitivity matrix-assisted laser desorption ionization mass spectrometry. J. Mass Spectrom. *30*:1768-1771.

We would like to thank Dr. R. Rush for his initial suggestion for the use of these columns and reading of the manuscript and Drs. Hsieng Lu and Tony Polverino for critical reading of the manuscript. We also wish to thank Scott Lauren for amino acid analyses. *Address correspondence to Scott D. Patterson, Amgen, Amgen Center, Mail Stop 14-2-E, 1840 DeHavilland Drive, Thousand Oaks, CA 91320-1789, USA. Internet: spatters@amgen.com*

Received 1 May 1996; accepted 26 July 1996.

Paul L. Courchesne and Scott D. Patterson
Amgen
Thousand Oaks, CA, USA

Long-Inverse PCR To Generate Regional Peptide Libraries By Codon Mutagenesis

BioTechniques 22:250-254 (February 1997)

We have developed a method to construct protein libraries by codon mutagenesis, which can be applied to any plasmid-encoded gene without the need for appropriate restriction sites and subcloning. The procedure entails synthesis of two oligonucleotides, in one of which selected codons are partially randomized, alkaline denaturation of template plasmid followed by long-inverse polymerase chain reaction (PCR) and treatment of DNA termini for blunt-end self-ligation to generate a mutagenized plasmid library.

Changes of amino acids by mutagenesis of cloned genes is increasingly used to examine structure-function relationships in proteins. Though useful, site-directed oligonucleotide mutagenesis is labor-intensive when many changes are required. Random mutagenesis of plasmids with chemical agents or *E. coli* mutator strains generates mutations outside of the gene of interest and becomes increasingly impractical as the size ratio of the plasmid to target gene increases. Procedures in which the target gene is amplified by PCR under conditions that increase the mis-incorporation rate by *Taq* DNA polymerase narrow the mutagenized region to the area of interest, but require high-efficiency cloning of the product and suitable restriction sites in both the product and target vector to generate a library of sufficient size (10). In addition, these methods are biased for transition mutations over transversions, which limits the amino acid substitutions at any given position. Furthermore, the genetic code itself minimizes the harmful effects of mutations (9). For example, pyrimidines in the second codon position specify predominately hydrophobic amino acids. An additional limitation in obtaining a broad spectrum of amino acid substitutions at a given codon is the extremely rare occurrence of an adjacent double-point mutation.

Codon-based mutagenesis or region-

al codon randomization generates highly complex yet condensed protein libraries (2,4). This procedure relies on complex oligonucleotide synthesis in which a portion of the appropriate codons are randomized using all 64 codons or a subset of codons on one DNA synthesis column while the wild-type codon is synthesized on another column. The matrices from the two columns are mixed and split again for the next randomized codon, and the process is repeated for each randomized codon. The mutation frequency is set by the ratio of resin in the column containing the mutated codon vs. that containing the wild-type codon and is biased to the wild-type sequence. Typically, conditions are chosen such that for a 20-amino acid sequence, 3000 clones will contain all possible single amino acid substitutions in addition to multiple substitutions at more than one position (2). The technique is not yet in widespread use, partly because of difficulties in generating a representative library and the necessity of introducing specific restriction enzyme sites.

We were recently unable to isolate temperature-sensitive mutations in the essential *Saccharomyces cerevisiae* gene QSR1 (11) by random mutagenesis procedures, which were successful in our hands with a gene of similar size (5). The possibility exists that a temperature-sensitive QSR1 allele was not obtained since not all proteins can be rendered thermolabile by the limited amino acid substitutions possible by single mutations (7). We thus developed a simple method of regional codon randomization mutagenesis. We targeted the mutagenesis to a highly charged 20-amino acid sequence on the premise that clusters of charged residues are likely to be on the surface of the protein and may be important for protein-protein interactions (12). With this procedure, we obtained 25 unique temperature-sensitive mutants from a single library.

The procedure is outlined in Figure 1 and simplifies randomization or site-directed mutagenesis of any region on any sized plasmid without the need for restriction enzymes. Long PCR (1) of two primers back to back (inverse PCR) (8) forms an 8-kb fragment of the entire plasmid. Long PCR allows for high-yield amplification of fragments in the 6–25-kb range and has the added benefit that fidelity is increased several fold because of the addition of a proofreading thermostable DNA polymerase. The fragment is treated with T4 DNA polymerase to remove 3′ overhanging A's added by *Taq* polymerase, then phosphorylated and self-ligated to regenerate a plasmid library incorporating the oligonucleotide mixture.

Three micrograms of the 8-kb plasmid pDEQ2 were denatured in 200 µL of 0.2 M NaOH and 2 mM EDTA for 10 min at 37°C, then precipitated at -20°C by adding 20 µL of 3 M Na acetate (pH 5.2) and 700 µL of ethanol. The pellet was washed and resuspended in 300 µL of water at a final concentration of 10 ng/µL. Alkaline denaturation before PCR increases product yield and eliminates unwanted background transformants from the template plasmid (3).

The Expand™ Long-Template PCR system (Boehringer Mannheim, Indianapolis, IN, USA) was used to generate a plasmid library incorporating the 20-amino acid randomized codon region. A partially randomized 81-nucleotide (nt) oligonucleotide was synthesized (2). A titration of $MgCl_2$ concentration was carried out by using all 3 supplied 10× buffers, and buffer No. 2 (2.25 mM $MgCl_2$) was chosen on the basis of a high product yield. All

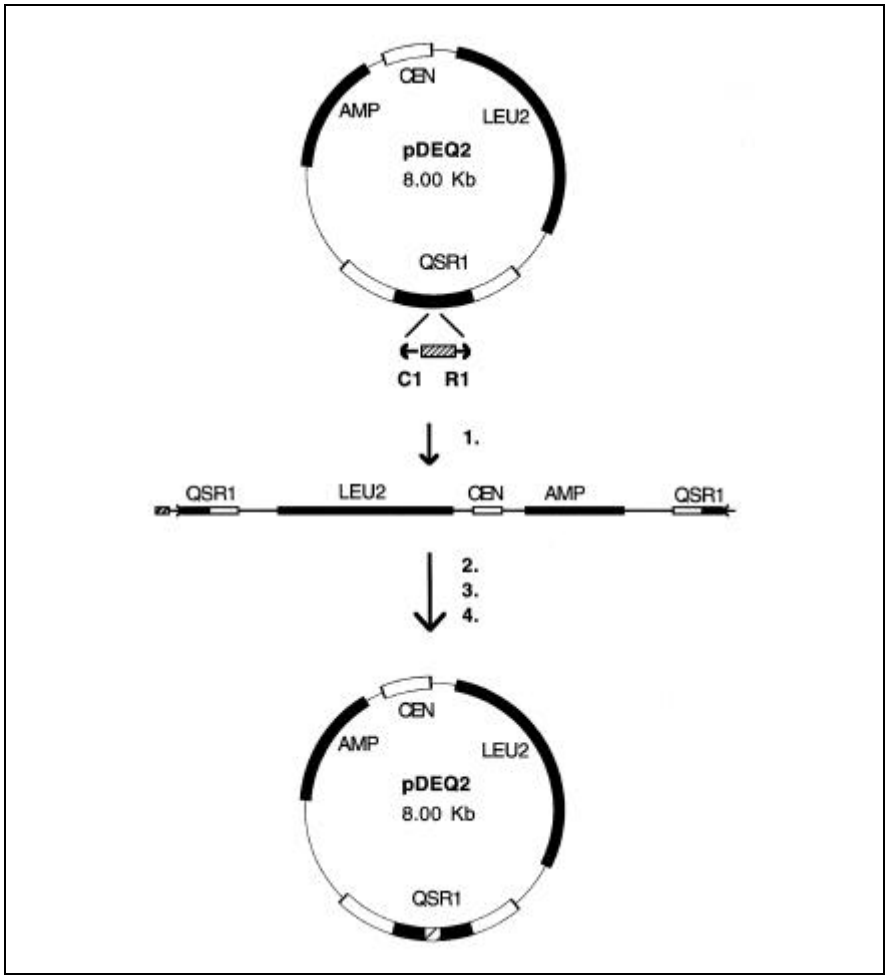

Figure 1. Mutagenesis by long-inverse PCR. Primers were designed in a back-to-back configuration (inverse PCR) such that self-ligation of the PCR product regenerates the template plasmid. The C1 primer is 21-nt long, the R1 is 81 nt in length with the 5′ 60 nt consisting of a mixture of wild-type and randomized codons (hatched area), while the 3′ 21 nt are wild-type for efficient annealing. The four steps are: 1. Long-inverse PCR with the two primers, 2. removal of 3′ overhanging A's with T4 DNA polymerase, 3. phosphorylation of 5′ termini and 4. self-ligation to regenerate a plasmid library incorporating the randomized oligonucleotide pool.

components were mixed on ice, and a 200-µL reaction was split into 4 thin-walled 0.6-mL PCR tubes (Robbins Scientific, Sunnyvale, CA, USA). Final concentrations were 350 µM dNTPs, 1× PCR buffer, 1.5 µL of enzyme mixture and 200 nM for each primer. After overlaying with mineral oil and placing in a thermal cycler preheated to 94°C, 0.2 ng of denatured template in 2 µL was added to each tube. Cycling conditions were 94°C for 10 s, 50°C for 2 min and 68°C for 6 min for 10 cycles followed by 15 cycles with a 20-s extension successively added to each 68°C extension step. At the end of 25 cycles, an additional 68°C extension for 10 min was included before cooling to 4°C. Depending upon the primer pairs, yields ranged from 1–10 µg of product.

The PCRs were pooled with 200 µL of 2× proteinase K buffer (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 2 mM EDTA, 1% sodium dodecyl sulfate) and 20 µg of proteinase K (Boehringer Mannheim) and incubated for 30 min at 37°C. After successive phenol, phenol/chloroform and chloroform extractions, the DNA product was precipitated with 1 mL of ethanol at -70°C for 10 min. After washing with 70% ethanol and air drying, the pellet was resuspended in water and blunted with 4 U of T4 DNA polymerase (Boehringer Mannheim) in the presence of 80 µM dNTPs for 2 h at 16°C. To change buffers, the DNA was purified by GENECLEAN® (BIO 101, Vista, CA, USA) and eluted in 16 µL of water. To this, 2 µL of 10× polynucleotide kinase buffer, 1 µL of 2 mM ATP and 1 µL (10 U) of polynucleotide kinase (Boehringer Mannheim) were added. After incubation at 37°C for 1 h, the 8-kb product band was gel-purified on a 0.7% agarose gel in 40 mM Tris acetate, 2 mM EDTA, pH 8.5, and eluted at a concentration of approximately 10 ng/µL, assuming a 50% loss with all manipulations. The kinasing step can be omitted if the primers are phosphorylated.

A self-ligation with 60 ng of DNA in a volume of 10 µL with 40 U of T4 ligase (New England Biolabs, Beverly, MA, USA) was carried out for 16 h at 16°C. Transformation competent DH5α cells were prepared by the high-efficiency procedure of Hanahan (6), and 8000 individual colonies were obtained per 5 µL of ligation transformed, which is greater than 10^5 clones per µg of DNA. A control lacking ligase gave a negligible background of 4 colonies. A library was made by pooling 18 000 colonies and purifying the plasmid DNA. Sequence analysis of 30 random clones revealed an average of 1.08 mutated codons per clone, whereas the 25 temperature-sensitive clones isolated had an average of 1.8 mutated codons per clone and were primarily single and double amino acid substitutions with a few triple substitutions. This higher rate of multiple mutated codons and the observation that only unique substitutions were obtained at a given codon reinforce the hypothesis that some genes are not rendered thermolabile by single-point mutations.

We have described a method of long-inverse PCR to make compact peptide libraries by codon mutagenesis on any plasmid. The procedure is also applicable to other gene modifications, and we have successfully used it for site-directed mutagenesis, protein truncations and epitope tagging.

REFERENCES

1. **Barnes, W.** 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. Proc. Natl. Acad. Sci. USA *91*:2216-2220.
2. **Cormack, B.P. and K. Struhl.** 1993. Regional codon randomization: defining a TATA-binding protein surface required for RNA polymerase III transcription. Science *262*: 244-248.
3. **Du, Z., D.A. Regier and R.C. Desrosiers.** 1995. Improved recombinant PCR mutagenesis procedure that uses alkaline-denatured plasmid template. BioTechniques *18*:376-378.
4. **Glasser, S.M., D.E. Yelton and W.D. Huse.** 1992. Antibody engineering by codon-based mutagenesis in a filamentous phage vector system. J. Immunol. *149*:3903-3913.
5. **Graham, L.A., U. Brandt, J.S. Sargent and B.L. Trumpower.** 1992. Mutational analysis of assembly and function of the iron-sulfur protein of the cytochrome bc1 complex in *Saccharomyces cerevisiae*. J. Bioenerg. Biomembr. *25*:245-257.
6. **Hanahan, D.** 1985. Techniques for transformation of *E. coli*, p. 115. *In* D.M. Glover (Ed.), DNA Cloning-A Practical Approach. Vol. 1. IRL Press, Oxford.
7. **Harris, S.D., J. Cheng, T. Pugh and J.R. Pringle.** 1992. Molecular analysis of *Saccharomyces cerevisiae* chromosome I: on the number of genes and the identification of essential genes using temperature-sensitive-lethal mutations. J. Mol. Biol. *225*:53-65.
8. **Hemsley, A., N. Arnheim, M.D. Toney, G. Cortopassi and D.J. Galas.** 1989. A simple method for site-directed mutagenesis using the polymerase chain reaction. Nucleic Acids Res. *16*:6545-6551.
9. **Lagerkvist, U.** 1981. Unorthodox codon reading and the evolution of the genetic code. Cell *23*:305-306.
10. **Leung, D.W., E. Chen and D.V. Goeddel.** 1989. A method for random mutagenesis of a defined segment using a modified polymerase chain reaction. Techniques *1*:11-15.
11. **Tron, T., M. Yang, F.A. Dick, M.E. Schmitt and B.L. Trumpower.** 1995. QSR1, an essential yeast gene with a genetic relationship to a subunit of the mitochondrial cytochrome bc1 complex, is homologous to a gene implicated in eukaryotic cell differentiation. J. Biol. Chem. *270*:9961-9970.
12. **Wertman, K.F., D.G. Drubin and D. Botstein.** 1992. Systematic mutational analysis of the yeast ACT1 gene. Genetics *132*:337-350.

This work was supported by Grant No. BE-204 from the American Cancer Society.
Address correspondence to Dominic P. Eisinger, Dartmouth Medical School, Vail 7200, Hanover, NH 03755, USA.

Received 3 June 1996; accepted 22 August 1996.

Dominic P. Eisinger and Bernard L. Trumpower
Dartmouth Medical School
Hanover, NH, USA

Family of Plasmid Vectors for the Expression of β-galactosidase Fusion Proteins in Eukaryotic Cells

BioTechniques 22:254-258 (February 1997)

*Lac*Z is a popular reporter gene for a variety of biochemical and genetic analyses. The enzymatic activity of its protein product, β-galactosidase, can be easily measured by well-developed spectrophotometric (4) and luminometric (1) assays in vitro and by histochemical staining in vivo (10). Moreover, it is possible to follow the spatial distribution of the gene product in situ, even in living animals, using lipophilic substrates (7,11). Antibodies against β-galactosidase made it possible to