Short Technical Reports

DNA Shuffling Method for Generating Estrogen Receptor α and β Chimeras in Yeast

BioTechniques 34:278-288 (February 2003)

ABSTRACT

To facilitate our study of the molecular basis for the estrogen receptor (ER) subtype selectivity of novel ligands, we used DNA shuffling to construct chimeric ERs having ligand binding domains derived from both ERα and ERβ. The efficiency of chimera generation was low with traditional DNA shuffling protocols. Furthermore, ER ligand binding domain sequences lack convenient restriction sites for introducing chimeric ligand binding domain sequences into expression vectors. To overcome these problems, we developed a modified strategy whereby chimeric sequences were exclusively amplified from among the reassambled products from DNA shuffling using a special pair of PCR primers whose 3′ ends specifically match the α and β sequences, respectively, and whose 5′ ends match sequences outside the ERβ ligand binding domain. When chimeric ligand binding domain DNA sequences, amplified with these primers, were co-transformed into yeast with a linearized expression vector for ERβ, an active expression vector was produced by homologous recombination. Twenty-two different crossover sites were found; most occurred when there was a stretch of eight or more identical base pairs in both sequences, and many were concentrated in the regions important for studying ligand binding and transactivation. This method should prove to be useful for generating chimeric gene products from parent templates that share relatively low sequence identity.

INTRODUCTION

Chimeric gene products can be produced from two homologous parent genes using restriction enzyme digestion and fragment ligation (15,19). However, the number and position of corresponding restriction sites are often very limited, making this strategy less general. Restriction sites can be engineered into the genes to facilitate gene fragment swapping, but this quickly becomes tedious when numerous chimeras are needed. Also, it is not always possible to introduce a restriction site at a desired position without changing the coded amino acids. Blunt-end ligation of PCR-generated fragments can be used to generate chimeras at any position, but this becomes cumbersome for more than one crossover and requires the rational design of crossover positions.

DNA shuffling involves in vitro or in vivo recombination methods for generating chimeric genes that rely on short stretches of identical DNA sequence in the homologous parent genes. DNA shuffling for chimeragenesis in gene families (13,17,18) has been widely used in directed evolution, examples of which include changing enzyme substrate specificity (24), improving enzyme thermostability (3), distinguishing functional and nonfunctional mutations (25), and probing protein structure-function relationships (4). The original Stemmer method for DNA shuffling uses DNase I fragmentation and gene reassembly (17); however, a variety of modified methods has been developed to improve the efficiency of in vitro recombination: staggered extension (5,26), shuffling with restriction enzyme-cleaved DNA fragments (6), shuffling using ssDNA (7,23), random chimeragenesis on transient templates (RACHITT) (1), the homology independent ITCHY method (12), and the

We would like to thank Adrienn Gabor for her exceptional assistance with software development and the design and construction of the incubator. This work was supported by National Institutes of Health grant nos. 5 F32 HL68457-02 to P.A.R. and 5P01 HL52813-08 to C.D.L., The G. Harold & Leila Y. Mathers Charitable Foundation (to C.D.L.), the Heartland Affiliate of the American Heart Association (Beginning Grant-in Aid 0160369Z to B.J.R.), and NATO-NSF DGE-0073179 to A.C. Address correspondence to Dr. Charles D. Little, Department of Anatomy and Cell Biology, University of Kansas Medical Center, 3901 Rainbow Blvd., Kansas City, KS 66160, USA. e-mail: clittle@kumc.edu

Received 30 August 2002; accepted 5 November 2002.

Paul A. Rupp1, Brenda J. Rongish1, Andras Czirik1,2, and Charles D. Little1
1University of Kansas Medical Center
Kansas City, KS, USA
2Eotvos University
Budapest, Hungary
SHIPREC method of Sieber et al. (16). These modified methods are often technically more challenging and require additional steps; some of them can only be applied to highly homologous genes. Some DNA shuffling methods have higher tolerance for parent genes with relatively low homology, but generally the efficiency of chimera generation decreases significantly as nucleotide sequence identities decline, and special approaches are needed.

We were motivated to generate a series of estrogen receptor (ER) chimeric constructs using DNA shuffling techniques to facilitate the study of the molecular basis of the subtype selectivity of some novel ER ligands. Since our targets, which are the ligand binding domains from ERα and ERβ share only 63% identity at the DNA level, they would be considered non-ideal substrates for application of the DNA shuffling technique. Here we describe the development of a simple but efficient and generally applicable DNA shuffling method to generate ER chimera constructs in a yeast system, with the sequences of their ligand binding domains derived from the two parent sequences in ERα and ERβ.

MATERIALS AND METHODS

Yeast Strain and Plasmid Constructs

The yeast reporter strain BJ-ECZ (ura3, trp1, leu2, pep4, pro1, prb1) containing nine tandem copies of an estrogen responsive LacZ reporter construct integrated into the ura3 locus was generated in our laboratory previously (22). To construct the human ERβ yeast expression vector, a 1.5-kb BamHI fragment of human ERβ cDNA from pNGV1-ERβ (10) was inserted into YEpERα (22) by replacing the 1.8-kb BamHI fragment containing human ERα cDNA.

DNA Shuffling

The strategy used is illustrated in Figure 1. It employs DNA shuffling (21) with several modifications to optimize the generation of crossovers and production of chimeras. Primers ERαA289f (5'-GCCAACCTTTGGC-
CAAGCC-3') and ERαS566r (5'-GCTTTGCTTCGTCTCCTCCAC-3') were used to amplify the cDNA of the ERα ligand binding domain from a plasmid template pCMV5-ERα. To amplify ERβ ligand binding domain cDNA, primers ERβA242f (5'-GGCCGCAAGGCAAGAAGTG-3') and ERβE516r (5'-CTCTGCGCGGTGAGCTC-3') were used in a PCR using pCMV5-ERβ (9) as a template. After gel purification using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA), 3 µg ERα and ERβ ligand binding domain were mixed, and 2.5 µL 1 M Tris-HCl (pH 7.5) and 2.5 µL 200 mM MnCl2 were added. The volume was brought to 49 µL with distilled water. The mixture was equilibrated at 15°C for 5 min. DNase I (1 U/µL; Invitrogen, Carlsbad, CA, USA) was freshly diluted 1:10 in distilled water, and 1 µL of it was added to the reaction. The digestion was performed at 15°C for 90 s and was stopped by adding 5 µL ice-cold stop buffer containing 50 mM EDTA and 30% (v/v) glycerol. The digested products were separated by electrophoresis in a 2% agarose gel. The DNA fragments were digested with restriction enzymes available at both ends of the synthetic minimal medium lacking tryptophan (SD-Trp). After 2–3 days at 30°C, the individual transformants were transferred to approximately 3 mL liquid SD-Trp medium and incubated at 30°C for two days. The plasmids were rescued from yeast cells using YEDER yeast DNA extraction reagent (Pierce Biotechnology, Rockford, IL, USA) and transformed into E. coli DH5α cells. The chimeric sequences were checked by DNA sequencing performed from plasmids purified from E. coli cells using BigDye™ and an ABI PRISM® 3700 sequencer (Applied Biosystems).

RESULTS AND DISCUSSION

We applied the DNA shuffling technique to generate ERs containing chimeric ligand binding domain sequences from the two ER subtypes, ERα and ERβ, to facilitate our studies of the molecular basis for the ER subtype-selectivity of novel, non-steroidal ligands. The nucleotide sequence identity between the two ER ligand binding domains is 63% when the alignment is specific for the amino acid sequence primers so that their 3' ends are matched only one and the other of the two parent genes, respectively. The PCR products are guaranteed to be chimeric, and wild-type sequences are completely eliminated (Figure 1C). Taking advantage of the plasmid construction by homologous recombination in yeast (8), we integrated the chimeric ligand binding domain sequences into the yeast expression vector YEpERβ through homologous recombination in yeast cells. We first amplified the two ER ligand binding domain cDNAs with an approximately 40-bp extension outside the ligand binding domain at both ends (approximately 0.8 kb) by standard PCR methods (Figure 2A). The purified PCR products underwent DNase I partial digestion, followed by separation on a 2% agarose gel (Figure 2B). After we isolated approximately 100-bp fragments from the agarose gel, we assembled them in a PCR without adding additional primers. The size of the assembled products ranged from approximately 300 bp to well over 1 kb. Since the latter length is longer than the input ligand binding domain sequence, this is an indication of the complexity of the reaction mixture (Figure 2C). We also tried to use approximately 50-bp fragments for the assembling reaction, but we failed to get the spectrum of the assembled products as shown in Figure 2C, most likely because of the low sequence identity between the two parent genes. After PCR amplification of this assembled product using primer pairs ERβαLBDf and ERβE516r or ERβA242f and ERβαLBDr (Figure 1B), a major product having the correct sequence size of the full-length ligand binding domain was obtained (Figure 2D). Because we had designed the two amplification primers so that their 3' ends were specific for one and the other of the two parent genes, respectively, the PCR products are guaranteed to be chimeric, and wild-type sequences are completely eliminated (Figure 1C). Taking advantage of the plasmid construction by homologous recombination in yeast (8), we integrated the chimeric ligand binding domain sequences into the yeast expression vector YEpERβ through homologous recombination in yeast cells.

We applied the DNA shuffling technique to generate ERs containing chimeric ligand binding domain sequences from the two ER subtypes, ERα and ERβ, to facilitate our studies of the molecular basis for the ER subtype-selectivity of novel, non-steroidal ligands. The nucleotide sequence identity between the two ER ligand binding domains is 63% when the alignment is specific for the amino acid sequence primers so that their 3' ends are matched only one and the other of the two parent genes, respectively. The PCR products are guaranteed to be chimeric, and wild-type sequences are completely eliminated (Figure 1C). Taking advantage of the plasmid construction by homologous recombination in yeast (8), we integrated the chimeric ligand binding domain sequences into the yeast expression vector YEpERβ through homologous recombination in yeast cells.
test strain BJ-ECZ (Figure 1D). The homologous sequences between the ends of the products and the linearized vector are at least 40 bp long, which is enough to elicit efficient recombination in yeast (11,14).

By functional assays, we found that the majority of the yeast transformants responded to the hormone 17β-estradiol in a β-galactosidase plate assay. By DNA sequencing, we found that about half of them had chimeric ligand binding domains with one crossover occurring at those sites indicated in Figure 3. A similar approach was used to generate ERα expression vectors containing chimeric ligand binding domains. Using this recombination-based method, we have also constructed ER expression vectors with their whole ligand binding domain swapped.

Throughout the approximately 0.75-kb ligand binding domain regions, 22 crossover positions have been found in the sequenced chimeric clones (Figure 3). Most of these sites are concentrated at both ends of the domain, corresponding to the regions having the highest sequence identity between the two ligand binding domains. The encoding residues from most of these sites are in the ligand binding pocket and are close to the ligand binding sites. These chimeric constructs have proved to be useful to us in studying the molecular basis of action of ER subtype-selective ligands through structure-function analyses of the activity of ERα or ERβ-specific ligands on ERα/ERβ chimeric proteins (J. Sun Ph.D. thesis, University of Illinois, Urbana, 2002; Reference 20; and manuscripts in preparation). We have found that a stretch of eight or more base pairs that are identical in both the ERα and ERβ sequences is sufficient to trigger an efficient recombination process in the shuffling reaction. In regions that share, overall, a very high DNA sequence similarity between the two ligand binding domains, as few as two identical base pairs sufficed to generate a crossover (Figure 3).

Parent sequence contamination is a significant problem associated with DNA shuffling. Here, in the PCR amplification step, we use primers that are specific to the two different ER subtype sequences so as to amplify only chimeric sequences out of the pool of reassembled products. Nevertheless, during our screen for ER expression vectors that contain shuffled ligand binding domain sequences, we found that half of the clones were wild type. Although we used double enzyme digestion to excise a small piece fragment in the wild-type ligand binding domain cDNAs, so as to make the DNA pool free from the complete wild-type sequence, some recombination did occur between the remaining ligand binding domain cDNA sequences in the plasmid and the amplified shuffled ligand binding domain cDNAs, and this gave a significant background of wild-type sequence (Figure 1E). One approach to reducing this background might be to remove most of the ligand binding domain sequence from the parent expression plasmid.

Theoretically, all of the amplified products should carry an odd number of crossovers. In reality, we found only a single crossover for each chimeric sequence; sequences containing three crossovers must be very rare, if they exist at all. This may be largely due to the annealing of DNA fragments derived from the same parental genes (homoduplex formation), whose probability is much higher than that of heteroduplex formation. By using the strategy illustrated in Figure 1F, which is a three-component homologous recombination system in yeast, we were able to generate chimeric ligand binding domains carrying two crossovers.

The strategies described here for generating chimeric gene products from the two homologous ER subtype genes provide a quick and efficacious way to perform DNA shuffling on tar-

![Figure 2. DNA shuffling between ERα and ERβ ligand binding domains.](image-url)
gets with low sequence identity that would not be good candidates for DNA shuffling by standard methodologies. This approach should be applicable to other homologous genes, including those for other nuclear hormone receptor subtypes.

REFERENCES


This work was supported by National Institutes of Health grant nos. PHS 5R01 CA18119 (B.S.K.) and PHS 5R37 DK13556 (J.A.K.), and American Cancer Society grant no. IRG99224001 (H.Z.). Address correspondence to Dr. Benita S. Katzenellenbogen, University of Illinois, Department of Molecular and Integrative Physiology, 524 Burrill Hall, 407 South Goodwin Avenue, Urbana, IL 61801-3704, USA. e-mail: katzenel@life.uiuc.edu

Received 26 September 2002; accepted 19 November 2002.

Jun Sun, John A. Katzenellenbogen, Huimin Zhao, and Benita S. Katzenellenbogen University of Illinois Urbana, IL, USA