Short Technical Reports

High-Fidelity In Vitro Recombination Using a Proofreading Polymerase

BioTechniques 30:530-536 (March 2001)

ABSTRACT

We describe a convenient PCR-based protocol for in vitro recombination of homologous genes, thereby minimizing the rate of associated point mutations. High-fidelity recombination conditions were obtained using Vent DNA polymerase, which, in contrast to Taq DNA polymerase, shows significant proofreading activity and ranges among the slowest thermostable DNA polymerases, allowing tight control of the polymerase-catalyzed DNA extension. To determine the mutagenesis rate and to analyze the efficiency of recombination, 89 clones

from a standard experiment were randomly selected for further analysis. Sequence comparison revealed that 21% (19/89) of the clones result from different recombination events in the marker-containing region (260 bp). The overall mutation rate is only 0.02%, which is the lowest rate thus far reported for in vitro recombination experiments.

INTRODUCTION

The natural evolution of proteins is a remarkably effective optimization process that is thought to be largely dependent on successive fixation of individual mutations and/or by recombination. Recently, several methods have been developed to mimic and accelerate nature's strategy to direct the evolution of protein function, in particular concerning ways to increase the genetic diversity in molecular populations. In

contrast to mutagenic PCR (2,12), error-prone reverse transcription of RNA (7), or directed mutagenesis (4,8), DNA shuffling (11,13,14,18) exchanges large functional domains of sequences in search for the best candidate molecules, thus accelerating the process of molecular evolution. Among the least elaborate methods for in vitro DNA recombination is the staggered extension process (StEP) method (18). StEP consists of priming the template sequences followed by repeated cycles of denaturation and extremely abbreviated annealing and polymerase-catalyzed extension. In each cycle, the growing fragments can anneal to different templates based on sequence complementary and extend further to create "recombination cassettes". However, the original protocol is based on *Taq* DNA polymerase, which does not exhibit proofreading activity and thus shows a rather high mutation frequen-

cy. Although these point mutations may provide useful diversity for some in vitro evolution applications, they are problematic for others. For example, much lower mutation rates are needed during the in vitro evolution of long genes or whole operons, particularly in structure-function studies of evolutionarily related sequences, in which the number of mutations is often correlated to inactivation of the protein(s) (1,15).

Here, we provide a convenient alternative protocol that randomly recombines genes with a very low rate of associated point mutagenesis. The protocol relies on the use of Vent DNA polymerase, which is known to exhibit proofreading activity. In addition, Vent DNA polymerase is also reported to have a low maximal extension rate of 1000 nucleotides/min and a processivity of about seven nucleotides/initiation event, as compared to 4000 nucleotides/min and 40 nucleotides/initiation

event, respectively, for *Taq* DNA polymerase (5). Therefore, Vent DNA polymerase is superior to *Taq* DNA polymerase when tight control of polymerase-catalyzed DNA extension is important (e.g., in StEP applications). This is readily demonstrated by the recombination of two genes encoding variants of *E. coli* penicillin acylase, including a thorough analysis of a large number of clones from a typical recombination experiment.

MATERIALS AND METHODS

Materials

Restriction enzymes and T4 DNA ligase were purchased from MBI Fermentas (St. Leon-Rot, Germany), and Vent DNA polymerase was purchased from New England Biolabs (Schwalbach, Germany). All oligonucleotides

and primers were obtained from Metabion GmbH (Martinsried, Germany). Plasmid preparations were performed using QIAprep® (Qiagen, Hilden, Germany). 6-nitro-3-(phenylacetamido)-benzoic acid (NIPAB) was from Sigma (Deisenhofen, Germany). PCR was performed in a UNO II Thermocycler™ (Biometra, Germany).

Plasmid Constructions

The DNA sequence coding for the wild-type N-terminus (A chain) of the penicillin acylase protein was amplified by PCR using primers PA-fwEco (5'-GCGAATTCATTAAAGAGGAGAA-ATTAACATGAAAAATAGAAATC-GTATG-3') and MCS-RV1 (5'-TTGG-ATACCCCGCATCGGCCCTTTG-GCCGGTCGGTCAAGCATTGGT-3') and plasmid pBT175 as template. The PCR product was digested with *Bbs*I and *Sfi*I and cloned into the *Bbs*I-*Sfi*I

Short Technical Reports

double-digested vector pQEB-MCS, generating plasmid pQEB-NPA. The *SfiI* restriction site was generated by introducing three silent mutations in the spacer region between the A and B chains (5'-CCGGCCAAAGGGGCC-3'). The presence of the new *SfiI* restriction site allows the construction of chimeric penicillin acylase genes using separately designed A and B chains. Based on the same pBT175 template, the DNA sequence encoding the C-terminus (B chain) was also amplified by PCR, this time using primers MCS-FW1b (5'-ACCAATGCTTGACCGA CCGGCCAAAGGGGCCGATG-GCGCACTGCTGGCG-3') and PA-RV3 (5'-CTT-GCAAGCTTAGTGAT-GGTGATGGTGTATGGCGCTGA-ACGTGCAACACTTCCT-3'). The PCR product was digested with *SfiI* and *HindIII* and cloned into the *SfiI/HindIII* double-digested vector pQEB-MCS, generating plasmid pQEB-CPA. For these reactions, 50 pmol each primer, 10 ng template DNA (pBT175), 0.2 mM each dNTPs, 1× Vent buffer, and 2.5 U Vent DNA polymerase were used. The PCR protocol consisted of an initial denaturation step for 5 min at 95°C and 25 cycles of 15 s at 95°C, 90 s at 68°C, and 30 s at 72°C.

The DNA sequence coding for the N-terminus of the penicillin acylase variant MT was designed to contain 12 point mutations as markers within the first 260 nucleotides (including the coding sequence of the signal peptide). Positions at which wild-type and variant MT differ from another are shown in bold type in Table 1. The plasmid harboring the MT sequence of the N-terminus of the penicillin acylase gene was constructed using the megaprimer method (9). The identity of all plasmid constructions was confirmed by automated DNA sequencing using an ABI 377 Prism® sequencer (Applied Biosystems, Foster City, CA, USA).

PCR-Mediated Homologous Recombination

Plasmids pQEB-NPA-MT and pQEB-NPA were prepared according to the miniprep protocols of Qiagen. The DNA concentration was determined spectrophotometrically. Equal amounts of pQEB-NPA-MT and pQEB-NPA

Table 1. DNA and Amino Acid Substitutions in the N-Terminal Part of the Penicillin Acylase Variant MT (Encoded by pQEB-NPA-MT) as Compared to the Wild-Type Sequence (Encoded by pQEB-NPA)

Marker Number	DNA		Amino Acid	
	Wild-Type	MT	Wild-Type	MT
1	GCT	GGT	Ala ₁₄	Gly₁₄
2	TGG	TGC	Trp ₂₀	Cys₂₀
3	GAG	GAA	Glu ₂₇	Glu ₂₇
4	AAG	AAC	Lys ₃₄	Asn₃₄
5	GGC	GGT	Gly ₄₁	Gly ₄₁
6	GCC	GCT	Ala ₄₇	Ala ₄₇
7	CTA	CTG	Leu ₅₃	Leu ₅₃
8	GTA	GTC	Val ₆₀	Val ₆₀
9	TTT	TTG	Phe ₆₇	Leu₆₇
10	CGC	CGT	Arg ₇₄	Arg ₇₄
11	GCG	GCT	Ala ₈₁	Ala ₈₁
12	TTT	TTC	Phe ₈₈	Phe ₈₈

Amino acid numbering starts with the N-terminal Met-codon of the signal peptide. DNA mutations and amino acid substitutions are shown in bold.

were used as templates in the subsequent PCR. Since the template length, primer binding sites, and GC content were the same for both templates, a similar amplification rate was assumed. For the PCR, 0.15 pmol each template, 50 pmol each primers PA-fwEco and MCS-RV1, 1× Vent buffer, 0.2 mM each dNTPs, and 2.5 U Vent DNA polymerase were used. The PCR protocol consisted of an initial denaturation step for 5 min at 95°C and 95 cycles of 1 s at 55°C and 30 s at 95°C.

Cloning and Analysis: Sequence and Enzyme Activity of Positive Clones

The PCR product was purified by agarose gel electrophoresis and followed by extraction using a QIAquick™ kit (Qiagen). The purified DNA was digested with *BbsI* and *SfiI* and cloned into the *BbsI-SfiI* double-digested vector pQEB-CPA. After transformation and selective plating using ampicillin, 89 randomly chosen clones carrying the A chain insert were sequenced using an ABI 377 sequencer. Proteins expressed from the clones were tested for enzyme activity using NIPAB as substrate, in 10 mM phos-

phate buffer, pH 7.5, at 37°C. Clones whose expressed proteins carried an active penicillin acylase were identified by the production of the yellow 6-nitro 3-aminobenzoic acid both by replica plating colonies on filter paper (16) as well as in liquid culture (6) using a robotic workstation (CyBi™-screen-machine; CyBio AG, Jena, Germany) including a POLARstar reader (BMG, Offenburg, Germany).

RESULTS AND DISCUSSION

In vitro recombination of DNA is a frequently used technique to recombine homologous DNA sequences during in vitro molecular evolution. However, most protocols introduce a considerable number of point mutations that could lead to significant fractions of inactive gene products, in particular when working with long genes. Here, we introduce a convenient in vitro recombination protocol with a very low point mutation rate. Two homologous types of DNA sequences coding for the N-terminus of penicillin acylase protein were used to test the recombination efficiency and the mutation rate of a modified StEP

Short Technical Reports

protocol (18). The two sequences, wild-type and MT, differ in 12 nucleotide positions (markers). Markers are distributed in such a way that, after every 20 bp wild-type sequence, one nucleotide is changed (Table 1).

An equimolar mixture of plasmid DNA, each containing the wild-type or MT genes, was subjected to a modified StEP recombination protocol using primers located 19 nucleotides before the first marker and 469 nucleotides after the last marker. The most important demand on efficient recombination by the StEP protocol is a stringent control of the extension reaction during PCR amplification. To limit polymerization (i.e., to stimulate recombination events), we used Vent DNA polymerase, which is known to have a low maximal extension rate and a low processivity of about seven nucleotides/

initiation event. In addition we reduced the extension times to 1 s. After 95 cycles, the PCR product was gel-purified, digested with the appropriate restriction enzymes, and ligated to the double-digested vector pQEB-CPA, thereby restoring the full-length gene of penicillin acylase. This gene library was amplified in *E. coli* XL1-Blue (Stratagene, La Jolla, CA, USA).

The recombination efficiency was analyzed by sequencing genes from 89 randomly selected clones. Among these, 19 clones were novel recombinants that differ from the parent genes, 43 clones were MT variants, and 27 clones were wild-type. Analysis of the recombinant molecules (Figure 1) showed that some recombinant clones are present in two or more copies. It was found that the clone with marker no. 1 was present twice (clones 43 and

145), the clone with markers no. 1–4 was present three times (clones 20, 92, and 134), the clone with markers no. 1–9 was present twice (clones 62 and 148), and the clone with markers no. 1–11 was present in two copies (clones 24 and 103). Strand switching occurred in all gene segments with approximately the same frequency (i.e., no statistically significant “hot spot” for recombination could be traced).

The frequency of occurrence of any particular marker from MT is in the range of 37%–68% (wild-type markers, 32%–63%). Since parent genes identified within the gene library also showed a slight predominance of MT variants over wild-type species, we tend to interpret the small total excess of MT markers as the consequence of small differences in the initial concentrations of the template DNA molecules in the StEP PCR, rather than assuming that template-switching events are not approximately random. All but three recombinants could most simply be explained to originate from a single strand switch (i.e., a single recombinant event). In contrast, the minimum number of crossovers required to generate clones 47, 104, and 161 was two. The rate of recombinants (19/89) appeared to be significantly smaller as compared to the original StEP protocol (18). However, this result can largely be attributed to the size of the analyzed gene segment, which was four times smaller in the experiment presented here. Indeed, the recombination rate—that is, the number of recombination events (22) per total number of analyzed bases [260 (marker region) $\times$ 89 (selected clones) = 23 140]—is comparable to recombination rates obtained with other in vitro recombination methods. The appearance of clones 104 and 161 gives a hint of the minimal distance between markers that could be recombined. In both cases, strand switching occurred at least twice within 40 or 60 bp, respectively. The presence of clones 43 and 145 shows that the strand extension can be tuned to the minimum of 21–36 nucleotides. Hence, the use of Vent DNA polymerase is favorable in terms of the efficient recombination of mutations separated by only a few base pairs. This feature of Vent DNA polymerase correlates well with its known processivity

Figure 1. Sequence analysis of the unscreened recombined gene library. White blocks represent wild-type (WT) sequences, and dark blocks represent mutant (MT) sequences. The total length of the region containing 12 marker mutations comprised 260 bp. Positions of new point mutations introduced during homologous recombination are represented by asterisks. Active (+) and inactive (-) variants were identified using standard penicillin acylase assays as described in the text.

Short Technical Reports

and polymerization rate (3,5).

Sequence analysis of the 89 randomly selected clones revealed no insertions or deletions. In a total of 38 500 sequenced bases (including PCR-amplified sequences flanking the marker-containing region), eight mutations were found, three of which appeared in recombinant clones. Thus, the overall mutation rate for this protocol is only 0.02% (8/38 500), which is three times less than the mutagenic rates of 0.05%–0.06% obtained with previous protocols (17,18). These results are in good agreement with the relative error rates measured for thermostable DNA polymerases (3). Compared to *Taq* DNA polymerase, the fidelity of Vent DNA polymerase is three times higher.

The current high-fidelity protocol has been successfully applied to distinguish between functional and nonfunctional mutations in the A chain of *E. coli* penicillin acylase. Among the amino acid substitutions introduced in variant MT, only the substitution Phe₆₇→Leu₆₇ is consistent with reduced enzyme activity. Furthermore, inactive clone 20 is similar to the active clones 92 and 134, except for a base mutation (GGC→GAC) that possibly is responsible for the reduced enzyme activity.

The protocol presented here shows the lowest mutation rate thus far reported for in vitro recombination experiments. At the same time, it proved to be nearly as efficient as previous methods in inducing recombination events. However, there might be cases when further improvement concerning the number of template switching events could be necessary. One way to achieve an increased recombination rate could be to further reduce the annealing-polymerization time (e.g., by using ultrafast thermocyclers) (10). At the current stage, however, the protocol already provides a powerful tool for applications concerning the molecular evolution of genes, operons, or even whole genomes.

REFERENCES

1. **Bowie, J.U., J. Reidhaar-Olson, W.A. Lim, and R.T. Sauer.** 1990. Deciphering the message in protein sequences: tolerance to amino acid substitutions. *Science* 247:1306-1310.
2. **Cadwell, R.C. and G.F. Joyce.** 1992. Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* 1:1-6.
3. **Cline, J., J.C. Braman, and H.H. Hogrefe.** 1996. PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerase. *Nucleic Acids Res.* 24:3546-3551.
4. **Horwitz, M.S.Z. and L.A. Loeb.** 1986. Promoters selected from random DNA sequences. *Proc. Natl. Acad. Sci. USA* 83:7405-7409.
5. **Kong, H., R.B. Kucera, and W.E. Jack.** 1993. Characterization of a DNA polymerase from the hyperthermophile archaea *Thermococcus litoralis*. Vent DNA polymerase, steady state kinetics, thermal stability, processivity, strand displacement, and exonuclease activities. *J. Biol. Chem.* 268:1965-1975.
6. **Kutzbach, C. and E. Ravenbush.** 1974. Preparation and general properties of crystalline penicillin-acylase from *Escherichia coli* ATCC 11105. *Hoppe-Seyler Z. Physiol.* 354:45-53.
7. **Lehtovaara, P.M., A.K. Koivula, J. Bamford, and J.K.C. Knowles.** 1988. A new method for random mutagenesis of complete genes: enzymatic generation of mutant libraries in vitro. *Protein Eng.* 2:63-68.
8. **Oliphant, A.R. and K. Struhl.** 1987. The use of random-sequence oligonucleotides for determining consensus sequences. *Methods Enzymol.* 155:568-582.
9. **Sarkar, G. and S.S. Sommer.** 1990. The "megaprimer" method of site-directed mutagenesis. *BioTechniques* 8:404-407.
10. **Schober, A., R. Günther, U. Tangen, G. Goldmann, T. Ederhof, A. Koltermann, A. Wienecke, A. Schwienhorst, and M. Eigen.** 1997. High-throughput screening by multi-channel glass fiber fluorimetry. *Rev. Sci. Instr.* 68:1-8.
11. **Shao, Z., H. Zhao, L. Giver, and F.H. Arnold.** 1997. Random-priming in vitro recombination: an effective tool for directed evolution. *Nucleic Acids Res.* 26:681-683.
12. **Spee, J.H., W.M. de Vos, and P. Kuipers.** 1993. Efficient random mutagenesis method with adjustable mutation frequency by use of PCR and dITP. *Nucleic Acids Res.* 21:777-778.
13. **Stemmer, W.P.** 1994. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* 91:10747-10751.
14. **Stemmer, W.P.C.** 1994. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* 370:389-391.
15. **Suzuki, M., F.C. Christians, B. Kim, A. Skandalis, M. Black, and L.A. Loeb.** 1996. Tolerance of different proteins for amino acid diversity. *Mol. Diversity* 2:111-118.
16. **Zhang, Q., L. Zhang, H. Han, and Y. Zhang.** 1986. A method for screening penicillin G acylase-producing bacteria by means of 2-nitro-5-phenylacetaminobenzoic acid test paper. *Anal. Biochem.* 156:413-416.
17. **Zhao, H. and F.H. Arnold.** 1997. Optimization of DNA shuffling for high fidelity recombination. *Nucleic Acids Res.* 25:1307-1308.
18. **Zhao, H., L. Giver, Z. Shao, J.A. Affholter, and F.H. Arnold.** 1998. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* 16:258-261.

This work was supported in part by grant BioFuture 0311852 from the Bundesministerium für Forschung und Technologie. We are grateful to the Genomics Laboratory, Georg-August University of Göttingen for their assistance concerning automated sequencing. Plasmid pBT175 was a kind gift of Dr. R. Müller (Boehringer Mannheim). Finally, the authors would like to thank Stephan Hettwer for helping us with our first StEPs. Address correspondence to Dr. Andreas Schwienhorst, Abteilung für Molekulare Genetik und Präparative Molekularbiologie, Institut für Mikrobiologie und Genetik, Grisebachstr. 8, 37077 Göttingen, Germany. e-mail: aschwie1@gwdg.de

Received 12 June 2000; accepted 20 September 2000.

Milena Ninkovic, Rüdiger Dietrich, Gülden Aral, and Andreas Schwienhorst
*Institut für Mikrobiologie und Genetik
Göttingen, Germany*