Gene, 164 (1995) 49–53
© 1995 Elsevier Science B.V. All rights reserved. 0378-1119/95/$09.50

GENE 09226

Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides

(Gene synthesis; plasmid synthesis; β-lactamase; in vitro recombination; assembly PCR; DNA shuffling; mutagenesis)

Willem P.C. Stemmer[a], Andreas Crameri[a], Kim D. Ha[b], Thomas M. Brennan[b] and Herbert L. Heyneker[b]

[a]*Affymax Research Institute, Palo Alto, CA 94304, USA; and* [b]*ProtoGene Laboratories, Inc., Mountain View, CA 94043, USA*

Received by F. Bolivar: 3 March 1995; Revised/Accepted: 15 June/23 June 1995; Received at publishers: 24 July 1995

SUMMARY

Here, we describe assembly PCR as a method for the synthesis of long DNA sequences from large numbers of oligodeoxyribonucleotides (oligos). The method, which is derived from DNA shuffling [Stemmer, Nature 370 (1994a) 389–391], does not rely on DNA ligase but instead relies on DNA polymerase to build increasingly longer DNA fragments during the assembly process. A 1.1-kb fragment containing the TEM-1 β-lactamase-encoding gene (*bla*) was assembled in a single reaction from a total of 56 oligos, each 40 nucleotides (nt) in length. The synthetic gene was PCR amplified and cloned in a vector containing the tetracycline-resistance gene (Tc^R) as the sole selectable marker. Without relying on ampicillin (Ap) selection, 76% of the Tc^R colonies were Ap^R, making this approach a general method for the rapid and cost-effective synthesis of any gene. We tested the range of assembly PCR by synthesizing, in a single reaction vessel containing 134 oligos, a high-molecular-mass multimeric form of a 2.7-kb plasmid containing the *bla* gene, the α-fragment of the *lacZ* gene and the pUC origin of replication. Digestion with a unique restriction enzyme, followed by ligation and transformation in *Escherichia coli*, yielded the correct plasmid. Assembly PCR is well suited for several in vitro mutagenesis strategies.

INTRODUCTION

The assembly of DNA sequences from oligos finds applications in DNA synthesis, gene expression and in vitro mutagenesis. Several publications describe the assembly of DNA sequences from oligos by ligation (Heyneker et al., 1976; Itakura et al., 1977; Goeddel et al., 1979), by the *Fok*I method of gene synthesis (Mandecki et al., 1988) and by self-priming PCR (Dillon and Rosen, 1990; Prodromou et al., 1992; Chen et al., 1994; Hayashi et al., 1994). Recently, DNA shuffling was introduced as a method for in vitro recombination by combinatorial assembly of genes from random fragments generated by partial DNaseI digestion (Stemmer, 1994a,b), or from a mixture of oligos and random fragments (Crameri and Stemmer, 1995). We describe here assembly PCR as a method for the single-step synthesis of DNA sequences up to 3 kb from large numbers of 40-nt oligos.

Correspondence to: Dr. W. Stemmer, Affymax Research Institute, 4001 Miranda Avenue, Palo Alto, CA 94304, USA. Tel. (1-415) 812-8727; Fax (1-415) 424-0832; e-mail: Pim_Stemmer@qmgates.affymax.com
Oligo inquiries to: ProtoGene Laboratories, Inc., 1615 Plymouth Street, Mountain View, CA 94043, USA. Tel. (1-415) 428-3810; Fax (1-415) 428-1600; e-mail: protogen@netcom.com

Abbreviations: aa, amino acid(s); Ap, ampicillin; bp, base pair(s); βGal, β-galactosidase; *bla*, gene encoding TEM-1 β-lactamase; dNTP, deoxyribonucleotide triphosphate; kb, kilobase(s) or 1000 bp; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; *ori*, origin of DNA replication; PCR, polymerase chain reaction; [R], resistant/resistance; Tc, tetracycline; u, unit(s); wt, wild type.

SSDI 0378-1119(95)00511-0

EXPERIMENTAL AND DISCUSSION

(a) Synthesis, shuffling and characterization of DNA coding for TEM-1 β-lactamase

We synthesized 56 oligos, 40 nt in length, which collectively encode both strands of the promoter region and the structural *bla* gene. The plus strand as well as the minus strand consists of 28 oligos configured in such a way that, upon assembly, complementary oligos will overlap by 20 nt (Fig. 1). We introduced five point mutations, creating five new restriction sites, to differentiate the synthetic gene from the naturally occurring *bla* gene in pBR322 (Bolivar et al., 1977; Watson, 1988). In our design of the synthetic gene, we incorporated *Sfi*I recognition sites flanking *bla* to facilitate cloning in pUC322-Sfi. Plasmid pUC322-Sfi is derived from pUC182Sfi (Stemmer, 1994a), by ligating an *Aat*II-*Msc*I fragment from pBR322 containing the Tc^R gene into pUC182Sfi digested with *Aat*II + *Sma*I.

The assembly PCR protocol consists of four steps: oligo synthesis, gene assembly, gene amplification and cloning. Since single-stranded ends of complementary DNA fragments are filled-in during the gene assembly process, cycling with DNA polymerase results in the formation of increasingly larger DNA fragments until the full-length gene is obtained. Note that DNA ligase is not being used in the assembly process. Experimental details of the protocol are outlined in the legend to Fig. 1.

The 1.1-kb PCR product was obtained as shown by 1% agarose gel electrophoresis (Fig. 2). After digestion with *Sfi*I, the expected 0.9-kb fragment was gel-purified and ligated into the vector pUC322-Sfi. *E. coli* K-12 transformants were selected on Tc plates. Of the Tc^R colonies, 76% (78/102) were resistant to 100 μg Ap/ml. Plasmids obtained from six Ap^R colonies were analyzed by restriction enzyme digestion. All plasmids contained the five restriction sites characteristic for the synthetic gene construct. The *bla* gene from one of these plasmids was sequenced: three point mutations were found, presumably introduced during the PCR amplification process.

(b) Synthesis, PCR assembly and characterization of DNA encoding a synthetic version of plasmid p182Sfi

We synthesized 132 oligos, 40 nt in length, as well as one 47-mer and one 56-mer, which collectively encode both strands of a synthetic version of the 2703-bp pUC182Sfi plasmid (Stemmer, 1994a). This plasmid is similar to pUC18 except for the introduction of two *Sfi*I sites flanking the *bla* gene. The overlap of complementary oligos is 20 nt. The oligos covering the *bla* gene were described in section *a*. A 47- and 56-mer were employed to complete the circle leaving two small single-stranded

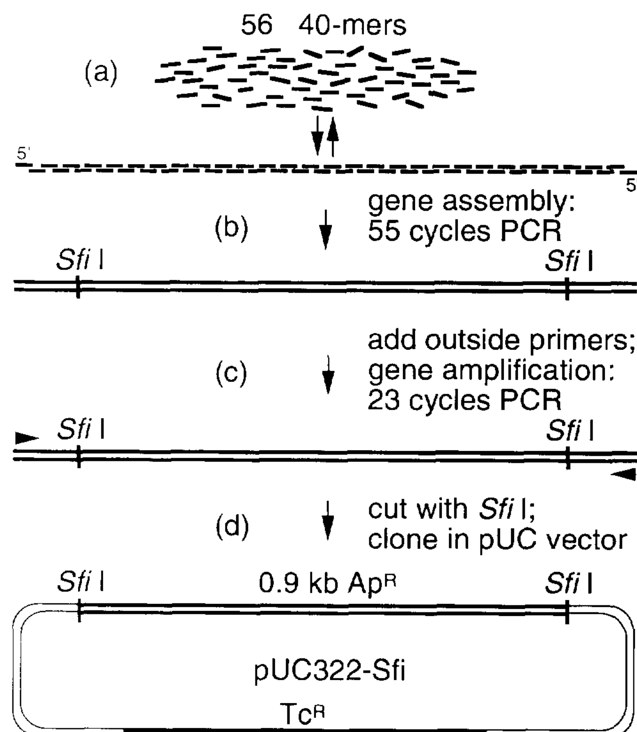

Fig. 1. Protocol for *bla* gene synthesis by oligo shuffling. (**a**) Oligo synthesis. An automated 96-well parallel-array oligonucleotide synthesizer (Lashkari et al., 1995) was used to prepare oligos on a scale of 50 nmol using standard phosphoramidite chemistry. After cleavage from the solid support (500 Å controlled pore glass) and deprotection with ammonia, the dried-down oligos were resuspended in distilled water at a concentration of approx. 250 μM and used without further purification. (**b**) Gene assembly. Equal volumes obtained from each of the 56 oligo solutions were combined and the mixture was subsequently diluted 100-fold in 20 μl PCR mix containing 10 mM Tris·HCl pH 9.0/2.2 mM $MgCl_2$/50 mM KCl/0.2 mM each dNTP/0.1% Triton X-100/1 u of *Taq* polymerase (Promega, Madison, WI, USA)/0.02 u of *Pfu* polymerase (Stratagene, La Jolla, CA, USA). The PCR program consisted of 55 cycles at 94°C for 30 s, 52°C for 30 s and 72°C for 30 s (MJ Research PTC-150 minicycler). (**c**) Gene amplification. Gene assembly reaction mixture (2.5 μl) was diluted 40-fold in 100 μl PCR mix containing 10 mM Tris·HCl pH 9.0/2.2 mM $MgCl_2$/50 mM KCl/0.2 mM each dNTP/0.1% Triton X-100/5 u of *Taq* polymerase (Promega)/0.1 u of *Pfu* polymerase (Stratagene)/2 outside primers at a concentration of 1 μM. The outside primers can be the same as the two oligos representing the 5' ends of the plus and minus strand. The PCR program consisted of 23 cycles at 94°C for 30 s, 50°C for 30 s and 72°C for 60 s. The DNA was extracted with phenol + chloroform, precipitated with ethanol and dissolved in 10 μl H_2O. (**d**) Cloning. The gene-amplified DNA was digested with *Sfi*I and the 0.9-kb fragment was purified from a 1% agarose gel after electrophoresis. Plasmid pUC322-Sfi was digested with *Sfi*I; the DNA fragment containing the *ori* and the Tc^R gene was also gel-purified. Two fragments were ligated with T4 DNA ligase and used to transform *E. coli* XL1-blue (Stratagene).

gaps, 7 and 16 nt in length, which will be filled-in during PCR. The 134 oligos were combined and assembled in a three-stage PCR protocol, resulting in a high-molecular-mass head to tail plasmid multimer. Experimental details are outlined in the legend to Fig. 3. Note that with this 'circular PCR' format (Stemmer, 1994b) there is no need

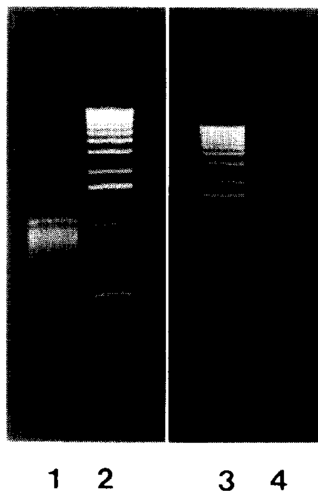

Fig. 2. Analysis of synthetic *bla* gene by 1% agarose gel electrophoresis. Lanes: **1**, 1.1-kb PCR product before digestion with *Sfi*I; **2** and **3**, 1-kb ladder; **4**, 0.9-kb PCR product after digestion with *Sfi*I.

for primer-mediated amplification. Digestion of the multimer with *Eco*RI, *Bam*HI, *Pst*I, *Bgl*II, *Sfi*I or *Nco*I resulted in the expected restriction pattern for the synthetic version of pUC182Sfi (Fig. 4). The linear fragment obtained after *Bam*HI digestion was gel-purified, treated with T4 DNA ligase and used to transform *E. coli* XL1-blue (Stratagene, La Jolla, CA, USA). A large number of blue colonies were obtained on IPTG + XGal + Ap plates. Minipreps of four colonies showed that all plasmids had the expected restriction digest pattern, including the five sites which were introduced in the *bla* gene (this paper).

(c) Conclusions

(*1*) The gene assembly method described here allows for rapid and cost-effective construction of long DNA sequences. It has been demonstrated (Stemmber, 1994a; Stemmer et al., unpublished data) previously that sequences of at least 10 kb, such as genes, gene libraries, plasmids, etc., can be reconstituted by DNA shuffling starting with 100 300-bp fragments obtained by nuclease digestion of the sequence of interest. Based on the data presented here, we expect that sequences in the range of 3 to 5 kb can be assembled from chemically synthesized oligos in a single reaction.

(*2*) The PCR assembly method facilitates several mutagenesis approaches, such as point mutagenesis (Stemmer, 1994a), combinatorial multiple cassette mutagenesis (Crameri and Stemmer, 1995) and doping, or mixing-in other nt during oligo synthesis (Hermes et al., 1989). Deliberate modifications to the nt sequence, for instance for rational design, can be made simply by substituting one or more new oligos followed by reassembly. To reduce the rate of PCR mutagenesis during assembly, the

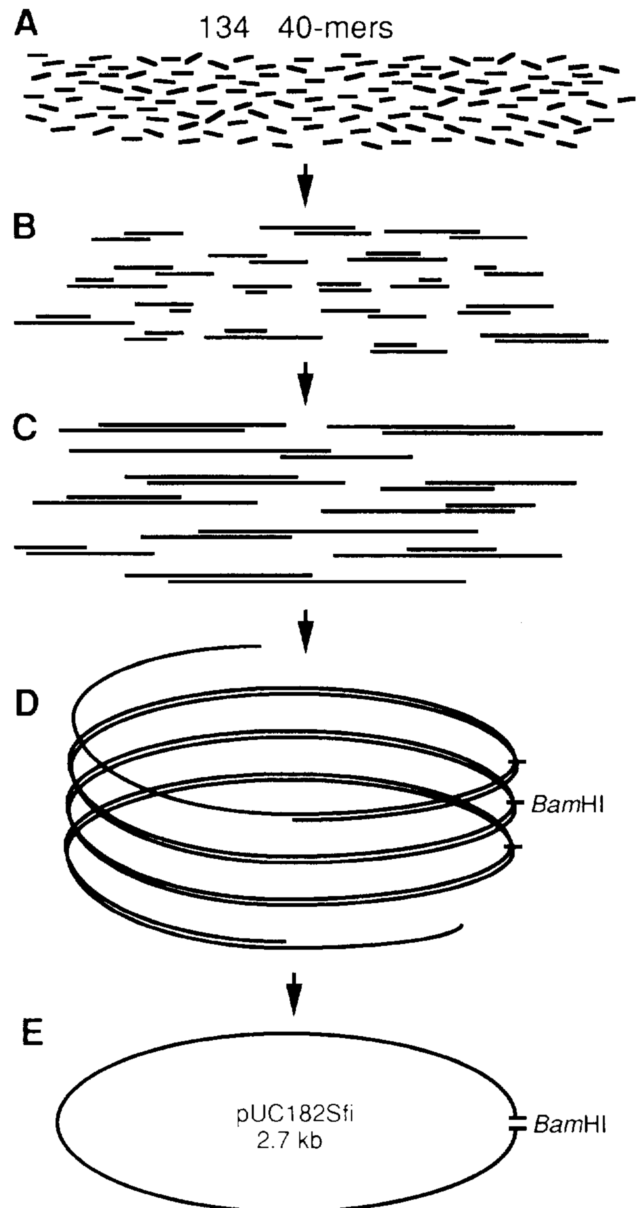

Fig. 3. Protocol for assembly of synthetic plasmid pUC182Sfi. Preparation of oligos was described in the legend to Fig. 1a. Equal volumes drawn from each of the 134 oligo solutions were combined to a final concentration of approx. 250 μM of mixed oligos, prior to a 250-fold dilution in 20 μl of Geneamp XL PCR mix (Perkin Elmer, Branchburg, NJ, USA). The amplification process consisted of three stages. The conditions for the first PCR program, which started with the pooled oligos (**A**), were: 40°C for 2 min, then addition of polymerase, 72°C for 10 s, then 40 cycles (94°C for 15 s, 40°C for 30 s and 72°C for 10 s – 1 s cycle). The reaction mixture was diluted three-fold with fresh PCR and polymerase mix. The conditions for the second PCR program (**B**) were: 25 cycles (94°C for 15 s, 40°C for 30 s and 72°C for 45 s + 1 s cycle). The reaction mixture was again diluted threefold in complete PCR mix. The conditions for the third PCR program (**C**) were: 20 cycles (94°C for 15 s, 40°C for 30 s and 72°C for 70 s + 1 s/cycle). The reaction products were analyzed by 1% agarose-gel electrophoresis (see Fig. 4). The high-molecular-mass assembled product (**D**) was digested with *Bam*HI (**E**) and the unit-length linear DNA was gel-purified, ligated and used to transform *E. coli* XL1-blue (Stratagene). It is important to use an *E. coli* strain which contains the *hsdR*17 mutation to prevent cleavage of the cloned DNA by the *Eco*K endonuclease.

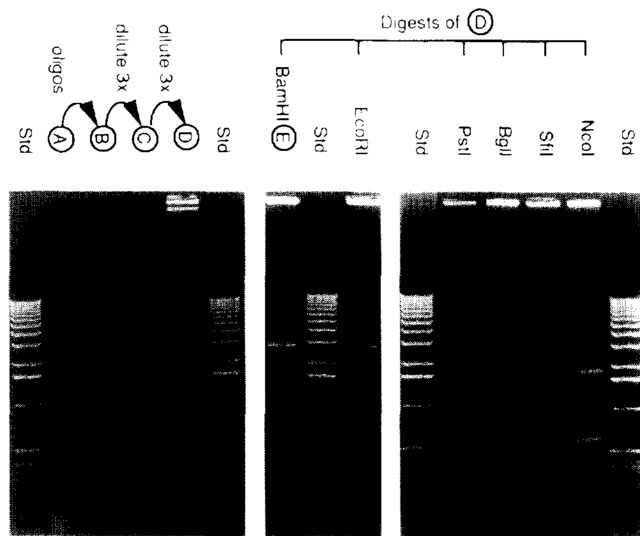

Fig. 4. Analysis of synthetic plasmid construct by 0.7% agarose gel electrophoresis. Lanes: Std, 1-kb ladder (Life Technologies); (A), equimolar mixture of the 134 oligos; (B), first stage assembly PCR product; (C), second stage assembly PCR product; (D), third stage assembly PCR product; (E), digest of product (D) with *Bam*HI; Other lanes, digest of product (D) with *Eco*RI, *Pst*I, *Bgl*I, *Sfi*I and *Nco*I.

addition of a proofreading polymerase may be important to ensure efficient long PCR reactions by combining high processivity with proofreading (Barnes, 1994).

(*3*) The experiments described in this paper were designed to demonstrate that the gene assembly method can be used for the construction of any gene as long as the gene product it codes for can be identified. For this reason we screened for Ap^R colonies rather than select for this trait. Since 76% of the colonies expressed a functional *bla* gene, we believe that in other gene assembly experiments, screening a limited number of colonies will be sufficient to identify the gene of interest.

(*4*) This is not the first time that an entire plasmid has been constructed from synthetic DNA building blocks. Mandecki et al. (1988; 1990) used the *Fok*I method for DNA synthesis of a pUC derivative, 2050 bp in length. Although elegant, the construction was cumbersome, requiring oligos up to 82 nt in length, which were purified, sequenced and cloned prior to assembly with DNA ligase.

Other investigators have reported on the construction of a synthetic gene without relying on DNA ligase for any part of the assembly process. Rosen et al. (1990) described the assembly by PCR of a 303-bp synthetic HIV-2 Rev gene starting with four 105-nt (!) oligos. Prodromou and Pearl (1992) described 'recursive PCR' as a novel technique for gene synthesis. Starting with ten oligos, 54–86 nt in length, and purified on polyacrylamide gels, they successfully assembled a 522-bp gene for human lysozyme. They gave special consideration to the overlapping regions, which were designed to give melting temperatures in the range of 52–56°C to ensure primer specificity and normalized annealing conditions.

(*5*) The real advantage of DNA shuffling as described in this paper lies in the robustness of the method: (*i*) all oligos used for gene synthesis were 40-mers, which can be synthesized efficiently without the need for further purification; (*ii*) we did not give special consideration to differences in melting temperature of the overlapping regions, which was fixed at 20 bp in length; (*iii*) the number of oligos that can be used in the one-step DNA assembly process is (surprisingly) high. We envision that larger DNA structures like bacteriophages and viruses can now be assembled from synthetic DNA fragments, either directly, or indirectly by ligation of a number of fragments obtained by assembly PCR.

Work is in progress on sequencing by hybridization and the development of methods to generate large arrays of oligonucleotides, in order to synthesize small quantities of oligos in large arrays on a glass surface using ink-jet technology for delivery of small amounts of chemicals to defined areas which are separated from each other by surface-tension barriers. This method will allow for the concomitant synthesis of many different oligos. Combined with assembly PCR, it will become possible in the near future to build very complex DNAs from off-the-shelf chemicals in a cost-effective way.

ACKNOWLEDGEMENTS

T.B. is supported by Department of Energy Grant No. DE FG03 93ER61615-A001 (1993).

REFERENCES

Barnes, W.M.: PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. Proc. Natl. Acad. Sci. USA 91 (1994) 2216–2220.

Bolivar, F., Greene, P.J., Betlach, M.C., Heyneker, H.L., Rodriguez, R. L., Boyer, H.W., Crosa, J.H. and Falkow, S.: Construction and characterization of new cloning vehicles, II. A multiple cloning system. Gene 2 (1977) 3987–3994.

Cadwell, R.C. and Joyce, G.F.: Randomization of genes by random PCR mutagenesis. PCR Methods Appl. 2 (1992) 28–33.

Chen, G.-Q., Choi, I., Ramachandran, B. and Gouaux, J.E.: Total gene synthesis: novel single step and convergent strategies applied to the construction of a 779 base pair bacteriorhodopsin gene. J. Am. Chem. Soc. 11 (1994) 8799–8800.

Crameri, A. and Stemmer, W.P.C.: Combinatorial multiple cassette mutagenesis creates all the permutations of mutant and wild-type cassettes: an antibody example. BioTechniques 18 (1995) 194–196.

Dillon, P.J. and Rosen, C.A.: A rapid method for the construction of synthetic genes using the polymerase chain reaction. BioTechniques 9 (1990) 298–300.

Goeddel, D.V., Kleid, D.G., Bolivar, F., Heyneker, H.L., Yansura, D.G., Crea, R., Hirose, T., Kraszewski, K., Itakura, K. and Riggs, A.D.: Expression in *Escherichia coli* of chemically synthesized genes for human insulin. Proc. Natl. Acad. Sci. USA 76 (1979) 106–110.

Hayashi, N., Welschof, M., Zewe, M., Braunagel, M., Dubel, S., Breitling, F. and Little, M.: Simultaneous mutagenesis of antibody CDR regions by overlap extension and PCR. Biotechniques 17 (1994) 310–314.

Heyneker, H.L., Shine, J., Goodman, H.M., Boyer, H.W., Rosenberg, J., Dickerson, R.E., Narang, S.A., Itakura, K., Lin, S. and Riggs, A.A.: Synthetic *lac* operator DNA is functional in vivo. Nature 263 (1976) 748–752.

Hermes, J.D., Parekh, S.M., Blacklow, S.C., Köster, H. and Knowles, J.R.: A reliable method for random mutagenesis: the generation of mutant libraries using spiked oligodeoxyribonucleotide primers. Gene 84 (1989) 143–151.

Itakura, K., Hirose, T., Crea, R., Riggs, A.D., Heyneker, H.L., Bolivar, F. and Boyer, H.W.: Expression in *E. coli* of a chemically synthesized gene for the hormone somatostatin. Science 198 (1977) 1056–1063.

Lashkari, D.A., Hunicke-Smith, S.P., Norgren, R.M., Davis, R.W. and Brennan, T.: An automated multiplex oligonucleotide synthesizer: development of high-throughput, low-cost DNA synthesis. Proc. Natl. Acad. Sci. USA 92 (1995) in press.

Mandecki, W. and Bolling, T.J.: *Fok*I method of gene synthesis. Gene 68 (1988) 101–107.

Mandecki, W., Hayden, M.A., Shallcross, M.A. and Stotland, E.: A totally synthetic plasmid for general cloning, gene expression and mutagenesis in *Escherichia coli*. Gene 94 (1992) 103–107.

Prodromou, C. and Pearl, L.H.: Recursive PCR: a novel technique for total gene synthesis. Protein Eng. 5 (1992) 827–829.

Stemmer, W.P.C.: Rapid evolution of a protein in vitro by DNA shuffling. Nature 370 (1994a) 389–391.

Stemmer, W.P.C.: DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proc. Natl. Acad. Sci. USA 91 (1994b) 10747–10751.

Watson, N.: A new revision of the sequence of plasmid pBR322. Gene 70 (1988) 399–403.