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\textsuperscript{a}, Andreas Crameri\textsuperscript{a}, Kim D. Ha\textsuperscript{b}, Thomas M. Brennan\textsuperscript{b} and Herbert L. Heyneker\textsuperscript{b}

\textsuperscript{a}Affymax Research Institute, Palo Alto, CA 94304, USA; and \textsuperscript{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 (\textit{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\textsuperscript{R}) as the sole selectable marker. Without relying on ampicillin (Ap) selection, 76% of the Tc\textsuperscript{R} colonies were Ap\textsuperscript{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 \textit{bla} gene, the α-fragment of the \textit{lacZ} gene and the pUC origin of replication. Digestion with a unique restriction enzyme, followed by ligation and transformation in \textit{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 \textit{FokI} 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.

Abbreviations: aa, amino acid(s); Ap, ampicillin; bp, base pair(s); \textit{βGal}, β-galactosidase; \textit{bla}, gene encoding TEM-1 β-lactamase; dNTP, deoxyribonucleotide triphosphate; kb, kilobase(s) or 1000 bp; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; \textit{ori}, origin of DNA replication; PCR, polymerase chain reaction; \textit{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 $\beta$-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 SfiI recognition sites flanking $bla$ to facilitate cloning in pUC322-Sfi. Plasmid pUC322-Sfi is derived from pUC182Sfi (Stemmer, 1994a), by ligating an AatII-MscI fragment from pBR322 containing the $Tc^R$ gene into pUC182Sfi digested with AatII + SmaI.

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 SfiI, 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 $\mu$g/ml.

(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 SfiI 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 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...
for primer-mediated amplification. Digestion of the multimer with *EcoRI*, *BamHI*, *PstI*, *BglII*, *SfiI* or *NcoI* resulted in the expected restriction pattern for the synthetic version of pUC182Sfi (Fig. 4). The linear fragment obtained after *BamHI* 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).

**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 *BamHI* (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 *hsdR17* mutation to prevent cleavage of the cloned DNA by the *EcoK* endonuclease.

(e) Conclusions

(1) The gene assembly method described here allows for rapid and cost-effective construction of long DNA sequences. It has been demonstrated (Stemmer, 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 (Cramer 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
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 ApR 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 FokI 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 tempera-


