

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Anuradha Subramanian Publications

Chemical and Biomolecular Research Papers --  
Faculty Authors Series

---

9-30-2007

## Rational *de novo* gene synthesis by rapid polymerase chain assembly (PCA) and expression of endothelial protein-C and thrombin receptor genes

Tarlan G. Mamedov  
*University of Nebraska - Lincoln*

Nisha V. Padhye  
*University of Nebraska - Lincoln*

Hendrik J. Viljoen  
*University of Nebraska - Lincoln*, hviljoen1@unl.edu

Anuradha Subramanian  
*Department of chemical Engineering, University of Nebraska Lincoln.*, asubramanian2@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/cbmesubramanian>

 Part of the [Chemical Engineering Commons](#)

---

Mamedov, Tarlan G.; Padhye, Nisha V.; Viljoen, Hendrik J.; and Subramanian, Anuradha, "Rational *de novo* gene synthesis by rapid polymerase chain assembly (PCA) and expression of endothelial protein-C and thrombin receptor genes" (2007). *Anuradha Subramanian Publications*. 3.

<https://digitalcommons.unl.edu/cbmesubramanian/3>

This Article is brought to you for free and open access by the Chemical and Biomolecular Research Papers -- Faculty Authors Series at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Anuradha Subramanian Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# Rational *de novo* gene synthesis by rapid polymerase chain assembly (PCA) and expression of endothelial protein-C and thrombin receptor genes

Tarlan G. Mamedov<sup>a</sup>, Nisha V. Padhye<sup>a, b</sup>, Hendrik Viljoen<sup>a</sup>, and Anuradha Subramanian<sup>a, \*</sup>

<sup>a</sup> Chemical and Biomolecular Engineering, 820 N 16th Street, University of Nebraska–Lincoln, Lincoln, NE 68588-0643, USA

<sup>b</sup> Megabase Research Products, 2726 N 48th Street, Lincoln, NE 68504, USA

\* Corresponding author – tel. 402 472-3463; fax 402 472-6889.

## Abstract

The assembly of synthetic oligonucleotides into genes and genomes is an important methodology. Several methodologies for such synthesis have been developed, but they have two drawbacks: (1) the processes are slow and (2) the error frequencies are high (typically 1–3 errors/kb of DNA). Thermal damage is a major contributor to biosynthetic errors. In this paper, we elucidate the advantages of rapid gene synthesis by polymerase chain assembly (PCA) when used in combination with smart error control strategies. We used a high-speed thermocycler (PCRJet<sup>®</sup>) to effectively minimize thermal damage and to perform rapid assembly of synthetic oligonucleotides to construct two different genes: endothelial protein C receptor (EPCR) and endothelial cell thrombin receptor, thrombomodulin (TM). First, the intact EPCR gene (EPCR-1, 612 bp) and a mutant EPCR-2 (576 bp) that lacked 4 N-linked glycosylation sites were constructed from 35 and 33 oligonucleotides, respectively. Next, for direct error comparison, another longer gene, the 1548 bp TM gene was constructed from 87 oligonucleotides by both rapid and conventional PCA. The fidelity and accuracy of the synthetic genes generated in this manner were confirmed by sequencing. The combined steps of PCA and DNA amplification are completed in about 10 and 22 min for EPCR-1, 2 and TM genes, respectively with comparable low errors in the DNA sequence. Furthermore, we subcloned synthetic TM, EPCR-1, EPCR-2 and native EPCR-1 (amplified from cDNA) into a *Pichia pastoris* expression vector to evaluate the expression ability, and to compare them with the native gene. Here, we illustrate that the synthetic genes, assembled by rapid PCA, successfully directed the expression of functional proteins. And, importantly, the synthetic and the native genes expressed proteins with the same efficiency.

**Keywords:** PCRJet<sup>®</sup> thermocycler, high-speed polymerase chain assembly, thermal damage of DNA and control, error comparison, gene synthesis and expression

## 1. Introduction

The chemical synthesis of a 77 bp long yeast ala-tRNA gene was first described in 1970 (Agarwal *et al.*, 1970). Refinements in chemical techniques (e.g., the use of a combination of organic synthesis followed by enzymatic 5' phosphorylation and DNA ligase-catalyzed joining) facilitated the total synthesis of a 126 bp gene encoding the *E. coli* suIII suppressor tyr-tRNA (Khorana, 1979). By the mid-1980s, automated phosphoramidite synthesizers were developed (Alvarado-Urbina *et al.*, 1981; Hunkapiller *et al.*, 1984; and Engels and Uhlmann, 1989). Today, automated phosphoramidite synthesis machines routinely synthesize deoxyoligonucleotides 15 to 100 nucleotides long. In the classic gene assembly method (Khorana, 1979), synthetic DNA oligonucleotides are 5' phosphorylated using T4 kinase, annealed to form overlapping duplexes, and

then enzymatically joined using T4 DNA ligase. A variation of that method was used to synthesize a 5386 bp ΦX174 RFI DNA molecule in 2 weeks (Smith *et al.*, 2003) where 42-mer oligonucleotides were 5' phosphorylated, gel-purified, annealed, and ligated using Taq ligase at 55 °C. After overnight high temperature ligation, the DNA reaction product was amplified using the polymerase chain assembly (PCA) strategy described by Stemmer *et al.* (Stemmer *et al.*, 1995). Recently, green-fluorescence-protein (GFPuv) was assembled from synthetic oligonucleotides by applying both coincidence filtering and consensus shuffling protocols to reduce errors in resultant DNA populations (Binkowski *et al.*, 2005). In previously described PCA strategies (Barnett and Erfle, 1990; Ciccarelli *et al.*, 1991), several overlapping non-phosphorylated 5'OH oligonucleotides were annealed in a single reaction at concentrations of less than 0.1 μM. After 20–30 cycles of PCA am-

plification, the concentrations of the two outer primers were raised to  $\sim 1 \mu\text{M}$  and another 30–35 PCA cycles were carried out (Stemmer *et al.*, 1995). Although only a few full-length DNA molecules assemble during the first PCR stage, they are selectively amplified during the second stage. Using this technique, Stemmer *et al.* (1995) constructed an  $\sim 2700$  bp long synthetic plasmid from 134 different overlapping 5'OH oligonucleotides. Refinements to the Stemmer *et al.* (Stemmer *et al.*, 1995) protocol have been described by Mehta *et al.* (1997), Gao *et al.* (2004), Shevchuk *et al.* (2004), and Young and Dong (2004). Subsequently, double-stranded DNA molecules 1.2–20 kb have been assembled by PCA processes.

To our knowledge, the recent report of a 1200 bp synthetic DNA fragment constructed in 1 week is the fastest kilobase DNA synthesis to date, though a synthesis of a 5386 bp DNA molecule in 2 weeks has been documented, too (Smith *et al.*, 2003; Young and Dong, 2004). In another study, parallel processing to construct about 10 fragments of  $\sim 5$  kb in 2 weeks was proposed (Kodumal *et al.*, 2004). However, in all reported experiments using PCA-based synthesis, error rates of 2–5 errors/1000 bp have been reported. For example, 11 sequence mistakes per synthetic 5386 bp  $\Phi\text{X174}$  RFI molecule were observed elsewhere (Smith *et al.*, 2003). At an error rate of 2 errors/1000 bp (Smith *et al.*, 2003), a synthetic 48.5 kb phage  $\lambda$  DNA molecule would contain  $\sim 97$  errors, a molecule  $>100$  kb long would contain  $>200$  errors, and a chromosome-sized 1 Mb DNA molecule would contain  $\sim 2000$  errors. Therefore, a clear understanding of the sources of synthetic errors, and methods to control them, become an essential part of the biomolecular engineering approach to prepare synthetic genes. We recently proposed a quantitative model of error accumulation during PCR amplification (Rouillard *et al.*, 2004). The proposed quantitative model predicts the accumulation of errors over the course of a PCR cycle. Our model analysis shows thermal damage contributes significantly to the total number of detectable errors; therefore, consideration must be given to thermal management of the PCA process.

Few precautions have been taken to minimize DNA thermal damage during gene assembly experiments. Furthermore, the thermocyclers used during these experiments are slow because they are often used in combination with conservatively long thermocycling protocols. We hypothesize that high-fidelity synthesis of DNA molecules requires an awareness of bio-synthetic errors and a practical strategy to minimize these errors. Error control requires a fundamental understanding of (1) the relationship between PCR kinetics and errors that occur during DNA replication and (2) thermal management of the process to minimize high temperature exposure. We used a combination of a fast thermocycler (in which DNA spends little time at elevated temperature) with kinetically optimized DNA biochemistry to optimize gene assembly. Using fast PCRJet<sup>®</sup> thermocycling, not only is the PCA process faster, but unwanted thermal damage (A + G depurination and cytosine deamination) is effectively minimized.

As a case study, we selected the *de novo* synthesis of two genes associated with endothelial protein-C receptor (EPCR). A synthetic gene  $\sim 612$  bp, EPCR-1, which encodes all the ex-

tracellular domains of EPCR and EPCR-2, a synthetic gene  $\sim 576$  bp that encodes all the extracellular domains of EPCR but lacks the four N-glycosylation sites, was constructed *de novo*. For error comparison, another longer gene, thrombin receptor gene, TM was constructed by both rapid and conventional PCA. A synthetic TM gene  $\sim 1548$  bp with 6-His and a factor Xa cleavage site was designed to encode all the extracellular domains of TM but lacking transmembrane domain. The correctness of the assembled genes was confirmed by DNA restriction digest analysis, and the fidelity of the synthetic gene constructs assembled by the PCA process were authenticated by dideoxy sequence analysis. We further verified the functionality of our method by cloning and expression in a *Pichia pastoris* expression system.

## 2. Materials and methods

### 2.1. Oligonucleotide design and synthesis

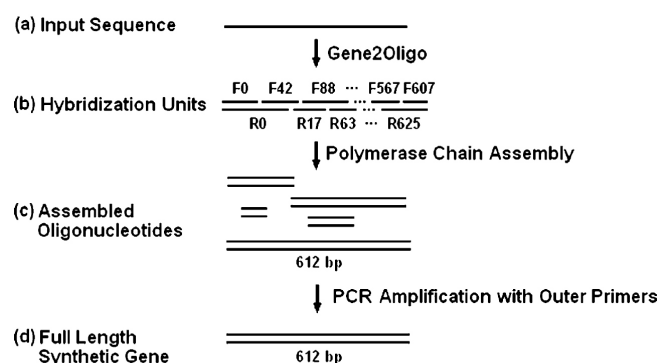
The template coding for EPCR-1 (i.e., residues 1 to 193) and EPCR-2 (i.e., residues 1 to 193 but lacking 36 bp corresponding to 4-N glycosylation sites) were based on the sequences of the EPCR. Oligonucleotides for a 612 bp EPCR-1 gene and a 576 bp mutant EPCR-2 gene were designed from a cDNA sequence of the human soluble EPCR deposited in the GenBank (GenBank accession no. L35545) using the Gene2Oligo (Rouillard *et al.*, 2004,) computer program. Using the same program, oligonucleotides for 1548 bp TM gene were designed from a cDNA sequence (accession no. BC035602) of the human TM after codon optimization using *P. pastoris* codons and adding 6-His tag and factor Xa site. The resulting oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). 5'OH oligonucleotides were synthesized by standard purification on a 25 nmol scale and stock solutions were obtained at  $50 \mu\text{M}$  in water. KOD Hot Start polymerase was obtained from Novagen (Madison, WI). A gel elute extraction kit was obtained from Qiagen (CA).

### 2.2. Instrumentation

PCA assembly was carried out in a PCRJet<sup>®</sup> thermocycler (Megabase Research Products, Lincoln, NE). The PCRJet<sup>®</sup> is a pneumatic device typically operated with 50–70 psig compressed air. Fast thermocycling with the PCRJet<sup>®</sup> is achieved by the flow of high velocity air through a thermostated reaction chamber that houses an array of capillaries. Two inlet air streams (one hot, one cold) are preconditioned to achieve the required temperature and flow rate in the reaction chamber (Moore, 2005; Quintanar and Nelson, 2002). Reactions are carried out in glass capillaries 25  $\mu\text{l}$  or less volumes.

### 2.3. Assembly PCR and amplification of EPCR-1, EPCR-2 and TM genes

The PCA (Stemmer *et al.*, 1995) of EPCR gene 1 was carried out in a 25  $\mu\text{l}$  volume containing  $0.1 \mu\text{M}$  of each oligonucleotide,  $200 \mu\text{M}$  of each dNTP, 4 mM  $\text{MgSO}_4$ ,  $400 \mu\text{g/ml}$



**Figure 1.** The assembly protocol for a 612 bp EPCR Gene 1. (A) 35 oligonucleotides were designed by the Gene2Oligo software. The length of overlapping hybridization units is optimized by the software to ensure specificity and uniform melting temperature. F and R correspond to the forward or (+) and the reverse or (–) strand of the input sequence. (B) Oligonucleotides are mixed at equimolar concentration and hybridized using polymerase chain assembly procedure (PCA). (C) PCA results in a few full length molecules and several incomplete molecules. (D) The full length synthetic gene is selectively amplified by PCR with outer primers.

non-acetylated BSA and 0.5 unit KOD Hot-Start Polymerase in 1× manufacturer's buffer. Assembly PCR reaction was conducted in a PCRJet® under the following conditions: a 30-s hot start at 92 °C followed by 30 cycles × (92 °C for 1 s and 68 °C for 4 s). Figure 1 explains the methodology to assemble the 35 oligonucleotides into 612 bp molecules. The full-length product from the assembly PCR reaction was amplified by PCR using the outer primers: 5'-ATCTCGAGAAAAGATTTGTAGCCAAGACGCCTC-3' (forward) and 5'-TAGCGGCCGCTTACGAAGTGTAGGACGGCTT-3' (reverse) at a concentration of 0.5 μM. PCR was carried out in 25 μl reaction volumes containing 0.5 μl assembly PCR mix, 200 μM of each dNTP, 4 mM MgSO<sub>4</sub>, 400 μg/ml non-acetylated BSA, 0.75 μM of each primer, and 0.5 unit KOD Hot-Start Polymerase in 1× manufacturer's buffer. PCR was conducted in the PCRJet® thermocycler under the following conditions: a 30-s hot start at 92 °C followed by 30 cycles × (92 °C for 1 s, 60 °C for 2 s, and 72 °C for 4 s).

For error comparison, a control EPCR was assembled under conventional PCR condition using the same PCRJet®. Assembly PCR reaction was conducted under the following conditions: a 30-s hot start at 92 °C followed by 30 cycles × (92 °C for 30 s and 68 °C for 60 s). PCR condition: a 30 s hot start at 92 °C followed by 30 cycles × (92 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s). Synthesis of EPCR-2 was undertaken in a manner similar to EPCR-1.

The PCA of TM gene was carried out in a 100 μl volume containing 0.1 μM of each oligonucleotide, 200 μM of each dNTP, 4 mM MgSO<sub>4</sub>, 400 μg/ml non-acetylated BSA and 0.5 units KOD Hot-Start Polymerase in 1× manufacturer's buffer. Assembly PCR reaction was conducted in a PCR-Jet® under the following conditions: a 30-s hot start at 94 °C followed by 30 cycles × (94 °C for 2 s and 60 °C for 10 s, 72 °C for 5 s). The full-length product from the PCA reaction was amplified by PCR using the following outer primers: 5'-ATCTCGAGAAAAGACATCATC-3' (forward) and

5'-TAGCGGCCGCTTAAGAATG-3' (reverse). PCR was carried out in 25 μl reaction volumes containing 0.25 μl assembly PCR mix, 200 μM of each dNTP, 4 mM MgSO<sub>4</sub>, 400 μg/ml non-acetylated BSA, 0.75 μM of each primer, and 0.5 units KOD Hot-Start Polymerase in 1× manufacturer's buffer. PCR was conducted in the PCRJet® thermocycler under the following conditions: a 30-s hot start at 94 °C followed by 35 cycles × (94 °C for 2 s, 60 °C for 3 s, and 72 °C for 10 s).

For error comparison, a control, TM gene was assembled under conventional PCR condition using the same PCR-Jet®. Assembly PCR reaction was conducted under the following conditions: a 30-s hot start at 94 °C followed by 30 cycles × (94 °C for 30 s and 60 °C for 60 s, 72 °C for 30 s). PCR amplification was conducted using above mentioned primers under the following conditions: a 30-s hot start at 94 °C followed by 35 cycles × (94 °C for 30 s and 60 °C for 30 s, 72 °C for 60 s).

After amplification, the purified PCR products from both rapid and conventional PCR were digested by Xho I and Not I and then directionally ligated with pPICZaA vector that was previously digested with both enzymes. The ligated plasmid was transformed into competent cells of *E. coli* strain TOP-10 (Invitrogen) and then positive 10 clones were selected from each sample. Purified plasmids from selected clones were sequenced at University of Nebraska–Lincoln sequencing facility using vector and gene specific primers.

#### 2.4. Vector constructions and expression of synthetic and native EPCR-1, EPCR-2 and TM genes in a *Pichia pastoris* system

Synthetic EPCR-1, EPCR-2, native EPCR-1 (corresponding to the residues 1–194, mature protein numbering) (Simmonds and Lane, 1999) and TM (corresponding to the residues 19–515 mature protein numbering, with 6 His and factor Xa site at the N-terminal) genes were expressed in *Pichia pastoris* (strain X-33) system (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol. In brief, cDNA encoding human soluble EPCR (GenBank accession no. L35545) was PCR amplified from the Gateway entry vector pENTR™221 (Invitrogen, Life Technologies) using primer pairs designed on the basis of the EPCR sequence (Simmonds and Lane, 1999). PCR was performed using Hot Start DNA polymerase (Stratagene). The respective sequences of the forward and reverse primers used to construct pPICZaA-native EPCR-1 were:

(forward) 5'-ATCTCGAGAAAAGATTTGTAGCCAAGACGCCTC-3'  
(reverse) 5'-TAGCGGCCGCTTACGAAGTGTAGGA-GC-GGCTT-3'.

The bold underlined nucleotide bases indicate restriction-enzyme digestion-site for Xho I and Not I, respectively.

A 612 bp EPCR-1, 576 bp EPCR-2 and 1548 bp TM genes were synthesized by adding Xho I and Not I restriction sites as described above. The amplified PCR product of EPCR-1 gene wild type and synthetic EPCR-1, EPCR-2 and TM genes



were digested by Xho I and Not I for at least 3 h at 37 °C and then directionally ligated with pPICZαA vector that was previously digested with both enzymes. The ligated plasmid was transformed into competent cells of *E. coli* strain TOP-10 (Invitrogen) and then positive clones were selected. The expression plasmids were isolated and sequenced in order to confirm accuracy of the synthetic genes and also to verify the ligations and inserts were correctly in-frame. Plasmids were linearized with Pme I and transformed into X-33 *Pichia pastoris* competent cells. Transformation, expression and methanol induction of recombinant proteins were performed according to the manufacturer's protocol provided for *Pichia* expression kit (Invitrogen). Samples of yeast culture supernatants were analyzed by SDS-PAGE and Western blotting. For evaluation of intracellular accumulation of TM, 0.25 g of the cells were disrupted with Mini-Beadbeater-48 using zirconia/silica beads and the cell lysates were centrifuged at  $10,000 \times g$  at 4 °C, the supernatant and pellet were analyzed by SDS-PAGE and Western blotting. N-glycosylation of the wild-type and synthetic EPCR-1 and TM were studied by digestion with endoglycosidase Endo H (New England Biolabs, Beverly, MA).

## 2.5. SDS-PAGE and Western blotting

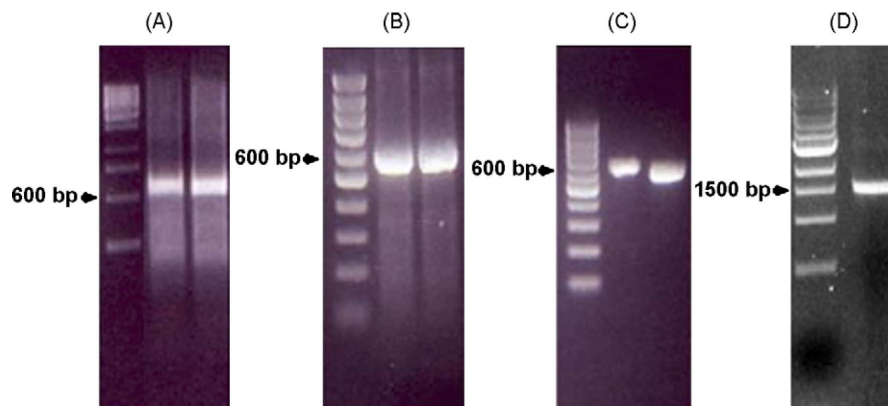
For SDS-PAGE, yeast culture medium containing expressed synthetic and native EPCR-1 proteins were concentrated 6-fold using Micron YM-30 (Millipore) and 20 µl of concentrated culture medium from each sample were separated in 10% Bis-Tris NuPAGE gels according to the manufacturer's protocol (Novex Pre-cast gel, Invitrogen). For Western blotting, the protein samples of 2 µl (native and synthetic EPCR-1) and 10 µl (EPCR-2) yeast culture medium and 10 µg of soluble protein (TM) were separated in 10% Bis-Tris, NuPAGE gels (Novex Pre-cast gel, Invitrogen) and the proteins were transferred to a PVDF membrane (Immobilon, Millipore) and then treated with a rabbit anti-human EPCR (for EPCRs) and anti His Tag Monoclonal Antibody (for TM). The signals were

visualized with ECL plus detection reagent and hyperfilm ECL from Amersham Biosciences (Piscataway, NJ).

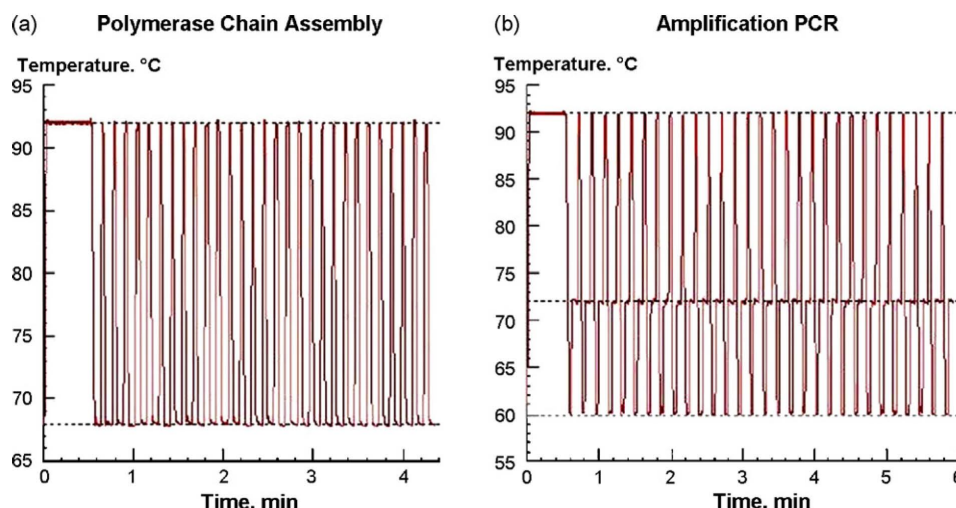
## 3. Results

### 3.1. Rapid polymerase chain assembly of EPCR-1, EPCR-2 and TM genes

The two-step polymerase chain assembly (PCA) method detailed by Stemmer *et al.* (1995) was adapted to high-speed assembly reaction conditions using the PCRJet®. EPCR-1 (Figure 2A), EPCR-2 (Figure 2B) and TM (Figure 2D) were successfully assembled from either 35, 33 or 87 synthetic oligonucleotides, respectively, using PCA followed by amplification of the full length desired gene by PCR. The gene assembly method we delineate here differs from previously described methods in a number of ways. Our method uses a fast pressurized gas thermocycler PCRJet® (Quintanar and Nelson, 2002) for rapid PCA. As a result, the combined steps of PCR assembly and DNA amplification are completed in about 10 min. The temperature versus time profile of the PCA step using the PCRJet® is shown in Figure 3a. The total time is about 4 min 30 s, including a 30-s hot start. In recent experiments, we reduced the number of cycles to 15 since only a few copies of the full-length assembled product are needed for further PCR amplification. The temperature versus time profile for the PCR amplification step is shown in Figure 3b. The average elongation rate of KOD Pol is 290–300 nucleotides per second (Griep *et al.*, 2006b); thus, a 4-s elongation is more than sufficient to complete extension. For error comparison, EPCR-1 and TM genes were also assembled by conventional PCR using the same PCRJet. 10 clones were selected from each sample and purified DNA was analyzed by sequencing. No DNA sequence errors were observed in the 612 bp EPCR-1 gene, synthesized by fast PCR, however 1.1 errors/kb observed for the same gene assembled by conventional PCR. Error rates of 0.53 errors/kb (fast PCA) versus 0.9 errors/kb (conventional PCA) were observed for TM gene. These collective findings



**Figure 2.** Two-step polymerase chain assembly of a 612 bp EPCR gene 1 (A) and a 576 bp EPCR-2 gene (B). (A) Eight microliters of each sample was analyzed on a 2% agarose gel. Lanes: (M) 1 Kb DNA Ladder (Fermentas Inc.); lane 1: 612 bp assembly PCR product with a 4-s elongation; lane 2: 612 bp assembly PCR product with 5-s elongation. (B) Eight microliters of each sample was analyzed on a 2% agarose gel. Lanes: (M) 100 bp DNA Ladder (Fermentas Inc.); lanes 1 and 2: 576 bp assembly PCR product with a 4-s elongation. (C) Gel purified EPCR genes 1 and 2. Eight microliters of each sample was analyzed on a 2% agarose gel. Lanes: (M) 100 bp DNA Ladder (Fermentas Inc.); lane 1: gel purified 612 bp EPCR gene 1, lane 2: gel purified 576 bp EPCR-2 gene. (D) Gel purified 1548 bp TM gene.



**Figure 3.** Temperature vs. time profiles for two-step polymerase chain assembly in the PCRJet<sup>®</sup> thermocycler. (a) First-round Assembly PCR: 30-s *KOD Pol* hot start at 92 °C followed by 30 cycles  $\times$  (1 s at 92 °C and 4 s at 68 °C) in a 25  $\mu$ l reaction volume. (b) Second round PCR amplification: 0.5  $\mu$ l of the first round reaction product was amplified in a new 25  $\mu$ l reaction volume containing 0.5  $\mu$ M outer primers: a 30-s hot start at 92 °C followed by 30 cycles  $\times$  (1 s at 92 °C, 2 s at 60 °C, and 4 s at 72 °C).

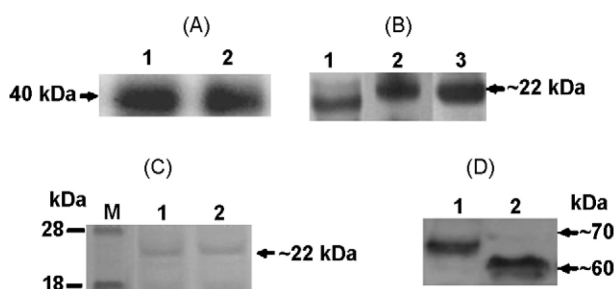
demonstrate that fast PCR eliminates thermal damage of DNA and reduces errors in DNA sequences.

### 3.2. Expression of synthesized genes in *Pichia pastoris*

The functionality of the synthesized genes was verified by expression of synthetic EPCR-1, EPCR-2, native EPCR-1 (native versus synthetic) and TM genes in the methylotrophic yeast *Pichia Pastoris*, a system highly successful for expressing heterologous proteins (Daly and Hearn, 2005). As a result, we were provided four constructs for recombinant protein expression: (1) pPICZaA/EPCR-1 (native), (2) pPICZaA/EPCR-1 (synthetic), (3) pPICZaA/EPCR-2 (synthetic) and pPICZaA/TM (synthetic). Regarding expression, all four constructs were linearized with Pme I at the vector 5'AOX1 promoter region, allowing for integration into the AOX1 locus of X-33 genome by homologous recombination. The transformants could be easily selected by their ability to grow on a Zeocin medium. The selected clones were induced according to the manufacturer's protocol provided for *Pichia* expression kit (Invitrogen) and significant secretions of EPCRs genes were achieved after 24 h, however, TM was not directed into the culture, but highly expressed intracellularly (see below). Yeast culture supernatants were analyzed by SDS-PAGE and Western blotting using rabbit anti-human soluble EPCR antibody.

The SDS-PAGE and Western blotting results presented in Figure 4 show proteins expressed by the native and synthetic EPCR-1 genes are similar with respect to their: (a) immunoreactivity with rabbit anti-human soluble EPCR antibody, (b) both showed the same migration pattern on a denaturing SDS-PAGE with a band of ~40 kDa (Figure 4A, lanes 1 and 2) and (c) incubation of both native and synthetic gene products by N-glycosidase F resulted in a unique and uniform lower band ~21 kDa (Figure 4B, lanes 2 and 3 and Figure 4C, lanes 2 and 3), indicating glycosylation was similar in both native and synthetic genes. Based on Western blotting (Figure 4A, lanes 1 and 2; Figure 4B, lanes 2 and 3) and SDS-PAGE

results (Figure 4C, lanes 1 and 2), the gene expression efficiency of both the native and the synthetic genes was found to be very similar. Figure 4B, lane-1 panel shows the EPCR-2 synthetic gene that lacked 4 glycosylation sites expressed in *Pichia pastoris* also cross-reacted well with rabbit anti-human soluble EPCR antibody and migrated as an ~20.5 kDa protein in the expected size. The expressed TM was analyzed by Western blotting using anti-His Tag Monoclonal antibody. As described above, highly disulfide bonded protein, recombinant TM, was not directed into the culture, however, using AOX1 promoter high level expression of rTM was observed



**Figure 4.** Expression of recombinant synthetic and native EPCR-1, EPCR-2 and TM genes in *P. pastoris*. (A) and (B) Native (A, lane 1), synthetic EPCR-1 (A, lane 2), and synthetic EPCR-2 (B, lane 1) were subjected to SDS-PAGE followed by Western blotting with a rabbit anti-human sEPCR antibody. 2  $\mu$ l (native and synthetic EPCR-1) and 10  $\mu$ l of yeast culture supernatant (EPCR-2) were electrophoresed under reducing condition. Lanes 2, 3 (B) show native and synthetic EPCR-1 after deglycosylation. Proteins from 2  $\mu$ l (native and synthetic EPCR-1) yeast culture supernatant were subjected to deglycosylation by digestion with Peptide-N-Glycosidase F endoglycosidase (New England, Biolabs). (C) Proteins from yeast culture medium were concentrated 6-fold using Micron YM-30 (Millipore) and 20  $\mu$ l of concentrated culture medium from each sample and subjected to 10% SDS-PAGE. (M) Protein standards; (1) *P. pastoris* clone that secretes native EPCR-1; (2) *P. pastoris* clone that secretes synthetic EPCR-1. (D) *P. pastoris* (harboring excreting TM) cell lysate was prepared as described in Section 2. 10  $\mu$ g of soluble proteins were run on SDS-PAGE followed by western blot using anti-His tag monoclonal antibody (lane 1) as described in Materials and Methods. 10  $\mu$ g of soluble proteins were treated with Endo H endoglycosidase (New England, Biolabs) and were subjected to deglycosylation (lane 2).

intracellularly as a ~70 kDa glycoprotein and a deglycosylated MM of ~60 kDa as expected size (Figure 4C). The molecular mass of expressed recombinant TM was higher under reducing condition (data not shown) when compared with non-reducing condition, similar to that of TM purified from human lung endothelial membrane preparations. Collectively, these findings provide evidence that (1) the synthetic genes synthesized by our method produced functional proteins and (2) high-speed polymerase chain assembly is accurate.

#### 4. Discussion

High-fidelity molecular DNA synthesis requires an awareness of biosynthetic errors and a practical strategy for the minimization of these errors. Error control requires a fundamental understanding of (1) the relationship between PCR kinetics and the errors that occur during polymerase-catalyzed nucleotide insertions and (2) thermal management to minimize high temperature exposure. The combination of a fast thermocycler (in which DNA spends very little time at elevated temperatures) with kinetically optimized DNA biochemistry is ideal to assemble synthetic DNA fragments.

Three sources of error may be identified during the construction of synthetic DNA molecules: (1) errors from the phosphoramidite synthesis of the oligodeoxyribonucleotides, (2) editing errors that occur during DNA polymerase-catalyzed enzymatic copying, and (3) errors that result from thermal damage to DNA. We used synthetic 5' hydroxylated oligonucleotides obtained from a commercial vendor; thus, deletions, mismatch errors, and/or guanine oxidation products that occur during phosphoramidite synthesis (23–25) were beyond our control. However, shorter synthetic oligomers have fewer errors – for that reason, we limited their size to  $\leq 40$  nucleotides.

The second source of error stems from enzyme editing mistakes when annealed oligomers are enzymatically extended. Extension errors depend on the fidelity of the polymerase. Extension errors can be reduced with a high fidelity polymerase enzyme and optimization of the biochemical reaction conditions during DNA extension. The fast thermostable  $\beta$ -type DNA polymerase *Pyrococcus kodakaraensis* (*KOD Pol*; Toyobo Co. Ltd., Osaka; Mizuguchi *et al.*, 1999) has proof-reading capability and it is preferred for rapid DNA amplification. The error frequency of *Pyrococcus kodakaraensis* (*KOD Pol*) under equimolar dNTP pool conditions was reported as 1.1 errors per million base pairs (Mizuguchi *et al.*, 1999). In other words, errors induced by *KOD Pol* are expected to occur at a rate significantly lower than the rate of errors stemming from phosphoramidite synthesis. A further advantage of *KOD Pol* is the high extension rate, typically 270–300 nucleotides per second (Griep *et al.*, 2006a; Griep *et al.*, 2006b), which further promotes rapid PCR assembly.

The greatest source of error in synthetic DNA molecule construction is thermally induced damage. Generally speaking, thermal damage to DNA is minimized when thermolabile nucleic acids spend as little time as possible at elevated temperatures ( $>50^\circ\text{C}$ ). For dsDNA, the rate of depurination increases four-fold for every  $10^\circ\text{C}$  increase in temper-

ature. The depurination rate of ssDNA is approximately 3.3 times higher than the depurination of dsDNA (Lindhahl and Nyberg, 1972). However, cytosine deamination of ssDNA occurs much faster than in dsDNA (Lindhahl and Nyberg, 1972). Therefore, enzymatic steps such as a  $65^\circ\text{C}$  overnight *Taq* ligase reaction (Smith *et al.*, 2003) are expected to result in substantial thermal depurination and C deamination to U in DNA on the order of 2–3 errors/1000 bp. Furthermore, many commonly used PCR thermocycling protocols, such as that of Saiki *et al.* (1985; 1 min at  $94^\circ\text{C}$ , 1 min at  $55^\circ\text{C}$ , and 1 min at  $72^\circ\text{C}$ ), spend unnecessarily long times at elevated temperatures. High-speed thermocycling protocols (Quintanar and Nelson, 2002; Whitney *et al.*, 2004; Wittwer *et al.*, 1990; Padhye *et al.*, 2002) that use parameters such as 0.5–2.0 s at  $94^\circ\text{C}$ , 0.5–10.0 s at  $55^\circ\text{C}$ , and 0.5–10.0 s at  $72^\circ\text{C}$  avoid excess “cooking” of nucleic acids, especially in the thermolabile ssDNA form. With fast PCRJet<sup>®</sup> thermocycling, not only is the PCR process faster, but unwanted thermal damage (A + G depurination and cytosine deamination) is effectively minimized. Our results demonstrated that fast PCR indeed eliminates thermal damage of DNA and significantly reduces errors in DNA sequences. No DNA sequence errors were observed in the 612 bp EPCR-1 gene, synthesized by fast PCR, however 1.1 errors/kb observed for the same gene assembled by conventional PCR using the same PCRJet. For TM gene, assembled by both fast and conventional PCR, error rates was 0.53 errors/kb and 0.9 errors/kb, respectively. As mentioned earlier, we have used synthetic 5' hydroxylated oligonucleotides obtained from a commercial vendor and any errors that were introduced during phosphoramidite synthesis were extrinsic. However, shorter synthetic oligomers are known to have fewer errors and for that reason, we limited their size to  $\leq 40$  nucleotides. In the case of EPCR-1 we did not observe any errors in the ten clones that were sequenced. It perhaps shows that the starting oligonucleotides that we used for PCA had an error rate of less than 1:6000. High-throughput gene synthesis on microchips using improved oligonucleotide synthesis method was recently reported (Tian *et al.*, 2004; Engels, 2005). Tian *et al.* (2004) developed a stringent hybridization method to remove error-containing oligonucleotides and using such oligos, a complete operon of 14.6 kb for 21 genes was assembled. The error rate for the oligonucleotides was estimated to be 1:7300 bp, confirming that error rate of synthetic oligonucleotides is one of the limiting step.

A quantitative model of error production during PCA process was recently developed (Pienaar *et al.*, 2006). Errors due to thermal damage stem mostly from three sources: (1) A + G depurination (Lindhahl and Nyberg, 1972), (2) oxidative damage of guanine to 8-oxoG (Frelon *et al.*, 2002 and Hsu *et al.*, 2004), and (3) cytosine deamination to uracil (Lindhahl and Nyberg, 1974). Oxidative damage can be reduced in PCR experiments by purging mixtures with argon to remove dissolved oxygen. Cytosine deamination is slowest at pH 8–8.5; the rate increases sharply at higher pH and more gradually at lower pH (when a small positive salt effect is also exhibited). The deamination rates are different for ssDNA and dsDNA; Fryxell and Zuckerkandl (2000) used data from three different



sources to obtain the following dsDNA deamination rate:  
C deamination of double stranded DNA

$$k_{\text{Cds}} = 2.66 \times 10^{10} e^{-32/RT} \text{ s}^{-1} \quad (R=0.00198 \text{ kcal/mol K}) \quad (1)$$

Lindahl and Nyberg (Lindahl and Nyberg, 1974) measured cytosine deamination rates in denatured *E. coli* DNA. At 95 °C and 80 °C, the rates are  $2.2 \times 10^{-7} \text{ s}^{-1}$  and  $1.3 \times 10^{-8} \text{ s}^{-1}$ , respectively. If an Arrhenius plot is fitted between the two points, the rate constant is:

C deamination of single stranded DNA

$$k_{\text{Css}} = 1.8 \times 10^{22} e^{-48.6/RT} \text{ s}^{-1} \quad (2)$$

Cytosine deamination of ssDNA does not only have a much larger activation energy than the deamination rate of dsDNA, but the reaction is much faster over the temperature ranges 60–95 °C.

Depurination reactions are strongly catalyzed at low pH. Based on the experimental data of Lindahl and Nyberg (1972), the rate constant for depurination of dsDNA at pH 7.4 is:

A + G depurination of double stranded DNA

$$k_{\text{DPds}} = 2.3 \times 10^{11} e^{-31/RT} \text{ s}^{-1} \quad (3)$$

The depurination rate constant of ssDNA is nominally 3.3 times higher than for dsDNA as illustrated elsewhere (28):

A + G depurination of single stranded DNA

$$r_{\text{DPss}} = 7.6 \times 10^{11} e^{-31/RT} \text{ s}^{-1} \quad (4)$$

To apply Equations (1), (2), (3) and (4) to a target template, it is necessary to know the degree of melting. At the beginning of a PCR cycle (i.e., the completion of primer/template annealing), the bulk of the template is in the single stranded form. As extension progresses, the fraction of the template that is double stranded increases. However, it is not correct to assume the polymerase/template complex marks the transition between dsDNA and ssDNA. Double stranded DNA may form bubbles, and localized melting may occur even at moderate temperatures. As the temperature increases to the denaturing temperature, the fraction of ssDNA template increases. Thus, a model that describes the helix/coil transitions is needed to assess the thermal damage.

A comparison of a protocol proposed by Saiki *et al.* (1985) (1 min at 94 °C, 1 min at 55 °C, and 1 min at 72 °C) with a typical PCRJet® protocol (0 s at 94 °C, 1 s at 55 °C, and 2 s at 72 °C) shows that total times per cycle (including ramp times) are 206 s and 4.66 s, respectively. First, the sample is heated

to 94 °C. No time is spent at the denaturation temperature and the sample is immediately cooled to 55 °C. It is annealed for 1 s and the temperature is raised at a rate of 47 °C/s to 72 °C, where it remains for 2 s before it is raised to 94 °C again for second cycle. The total time for one cycle is 4.7 s. The total thermal errors over one PCR cycle (depurination and C deamination) are 7.2/10<sup>6</sup> bp for the slow protocol and 0.026/10<sup>6</sup> bp for the fast protocol. The bulk of the damage occurs during the denaturation hold. Thus, a 50-fold reduction in the PCR period leads to a 270-fold reduction in errors.

#### 4.1. Rapid polymerase chain assembly of DNA with the PCRJet®

The two-step polymerase chain assembly method detailed by Stemmer *et al.* (1995) was adapted to high-speed assembly reaction conditions using the PCRJet® to assemble a synthetic 612 bp EPCR-1 and 1548 bp TM genes from 35 or 87 synthetic 5'OH oligonucleotides, respectively. A web-based tool (Gene2Oligo), which designs oligonucleotides of uniform melting temperature,  $T_m$ , was used to design oligonucleotides. For both EPCR-1 and EPCR-2, the  $T_m$  ranges from 67 °C to 75 °C, with the average being 69.8 °C. For TM gene the  $T_m$  ranged from 60 °C to 66 °C, with the average being 62 °C. Hence, we performed the assembly reactions at the highest possible temperature since the mismatches have a stronger destabilizing effect at high temperature – this diminishes the probability of incorporating mutated oligonucleotides into a final product. This full-length assembled product was further amplified by PCR. The total time needed for PCRJet® polymerase chain amplification of the 612 bp EPCR-1 gene starting with a mixture of 35 synthetic 5'OH oligonucleotides was about 10.0 min and for the 1548 bp TM gene was about 22 min. The fidelity of the synthesized PCA constructs was authenticated by dideoxy sequence analysis. The DNA sequence obtained was consistent with the cDNA sequence of the soluble EPCR and TM deposited in the GenBank (GenBank accession nos. L35545 and D00210). On an average 10 clones were sequenced for each PCA construct assembled and the types of observed mutations was shown in the Table 1. As we can see from the table errors stem mostly from (A + G) depurination, oxidative damage of G to 8-oxoG and C deamination. These results then support our hypothesis that thermal damage is a source of error in the gene assembly process. Additionally, the discrepancy between the fast protocol error rates for EPCR-1 (zero) compared to TM (0.53 errors/kb) may have arisen due to the sizes of TM (~1528 bp) versus EPCR (~612 bp).

**Table 1.** Comparison of observed mutations in the EPCR-1 (612 bp) and TM genes (1548 bp) assembled by conventional and fast PCR

Assembled genes	Number of clones sequenced	Errors/kb <sup>a</sup>	Total errors in 10 sequenced clones	Types of observed mutations (%)			
				A	T	C	G
EPCR-1 conventional PCR	10	1.1	7	28.5	14.28	42.85	14.28
EPCR-1 fast PCR	10	0.0	0	0.0	0.0	0.0	0.0
TM-1 conventional PCR	10	0.9	14	35.71	28.57	14.28	21.42
TM-1 fast PCR	10	0.53	8	25.0	0.0	50.0	25.0

<sup>a</sup> The numbers are reported in the manuscript.



Although, the final, optimal conditions are mentioned in this paper, the best annealing temperature for assembly was always determined by an optimization scheme, wherein the assembly process was started 8° to 10° below the average melting temperature, and the temperature was gradually increased to produce the highest yield. As such, we were able to achieve both specificity and rapid assembly. The time required for annealing during the assembly step is longer than what is required for standard PCR – this translates to the time required for the oligos to hybridize. The annealing time has been optimized with gene fragments of various lengths (data not shown); we used the standard time based on the number of oligos. After the assembly and expression of synthetic wild type was authenticated, the deletion mutant EPCR-2 (576 bp) that lacked 4 glycosylation sites was assembled from a template, resulting in no detectable errors in the DNA sequence of the assembled EPCR-2 gene product.

The following conclusions can be drawn from the work we elucidated here: (1) our *de novo* gene assembly methodology allows genes to be synthesized in an expedited time scale a once important proteins are identified; (2) the fast PCA process affords the efficient assembly of genes from templates of putative DNA sequences with comparable low errors; (3) the resulting synthetic genes can further be cloned and over-expressed to provide a large quantities of pure protein, which may then be used to generate high-affinity reagents in protein arrays, and (4) though we delineated the synthesis of a deletion mutant, our strategy can also be used to generate genes with substitutions, insertions, or targeted rearrangements in an expeditious manner. We believe the proposed strategy of rapid *de novo* gene synthesis—when coupled with proteomics—has the potential to accelerate the development of new disease markers and drug targets via the analysis of clinically relevant molecular events. Our future research efforts will aim to synthesize larger genes using an automated strategy in the rapid PCA thermocycler and also undertake the PCA and PCR of genes with a high CG content to generate error-time-temperature plots.

## Acknowledgments

The authors thank Casey Kotera for his technical assistance. This work was supported in part by funds from the NIH (1 R21 RR022860-01).

## References

- AGARWAL ET AL., 1970 — K.L. Agarwal, H. Buchi, M.H. Caruthers, N. Gupta, H.G. Khorana, K. Kleppe, A. Kumar, E. Ohtsuka, U.L. Rajbhandary, J.H. Van de Sande, V. Sgaramella, H. Weber and T. Yamada, Total synthesis of the gene for an alanine transfer ribonucleic acid from yeast, *Nature* **227** (1970), pp. 27–34.
- ALVARADO-URBINA ET AL., 1981 — G. Alvarado-Urbina, G.M. Sathe, W.-C. Liu, M.F. Gillen, P.D. Duck, R. Bender and K.K. Ogilvie, Automated synthesis of gene fragments, *Science* **214** (1981), pp. 270–273.
- BARNETT AND ERFLE, 1990 — R.W. Barnett and H. Erfle, Rapid generation of DNA fragments by PCR amplification of crude, synthetic oligonucleotides, *Nucleic Acids Res.* **18** (1990), p. 3094.
- BINKOWSKI ET AL., 2005 — B.F. Binkowski, K.E. Richmond, J. Kay-sen, R. Sussman and P.J. Belshaw, Correcting errors in synthetic DNA through consensus shuffling, *Nucleic Acids Res.* **33** (2005), p. e55.
- CICCARELLI ET AL., 1991 — R.B. Ciccarelli, P. Bunyuzlu, J. Huan, C. Scott and F. Oakes, Construction of synthetic genes using PCR after automated DNA synthesis of their entire top and bottom strands, *Nucleic Acids Res.* **19** (1991), pp. 6007–6013.
- DALY AND HEARN, 2005 — R. Daly and M.T. Hearn, Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production, *J. Mol. Recognit.* **18** (2005), pp. 119–138.
- ENGELS, 2005 — J.W. Engels, Gene Synthesis on Microchips, *Angew. Chem. Int. Ed.* **44** (2005), pp. 7166–7169.
- ENGELS AND UHLMANN, 1989 — J.W. Engels and E. Uhlmann, Gene Synthesis, *Angew. Chem. Int. Ed.* **28** (1989), pp. 716–733.
- FRELON ET AL., 2002 — S. Frelon, T. Douki and J. Cadet, Radical oxidation of the adenine moiety of nucleoside and DNA: 2-hydroxy-2'-deoxyadenosine is a minor decomposition product, *Free Radical Res.* **36** (2002), pp. 499–508.
- FRYXELL AND ZUCKERKANDL, 2000 — K.J. Fryxell and E. Zuckerkandl, Cytosine deamination plays a primary role in the evolution of mammalian isochores, *Mol. Biol. Evol.* **14** (2000), pp. 1371–1383.
- GAO ET AL., 2004 — X. Gao, P. Yo, A. Keith, T. Ragan and T. Harris, Thermodynamically balanced inside-out (TBIO) PCR-based gene synthesis: a novel method of primer design for high-fidelity assembly of longer gene sequences, *Nucleic Acids Res.* **31** (2004), pp. e143–e164.
- GRIEP ET AL., 2006a — M. Griep, C.A. Kotera, R.M. Nelson and H.J. Viljoen, Kinetics of the DNA Polymerase *Pyrococcus Kodakaraensis*, *Chem. Eng. Sci.* **61** (2006), pp. 3885–3892.
- GRIEP ET AL., 2006b — M. Griep, S. Whitney, R.M. Nelson and H.J. Viljoen, DNA Polymerase Chain Reaction: A Model of Error Frequencies and Extension Rates, *AIChE J.* **52** (2006), pp. 384–392.
- HSU ET AL., 2004 — G.W. Hsu, M. Ober, T. Carell and L.S. Beese, Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase, *Nature* **431** (2004), pp. 217–221.
- HUNKAPILLER ET AL., 1984 — M. Hunkapiller, S. Kent, M. Caruthers and L. Hood, A microchemical facility for the analysis and synthesis of genes and proteins, *Nature* **310** (1984), pp. 105–111.
- KHORANA, 1979 — H.G. Khorana, Total synthesis of a gene, *Science* **203** (1979), pp. 614–625.
- KODUMAL ET AL., 2004 — S.J. Kodumal, K.G. Patel, R. Reid, H.G. Menzella, M. Welch and D.V. Santi, Total synthesis of long DNA sequences: synthesis of a contiguous 32-kb polyketide synthase gene cluster, *Proc. Natl. Acad. Sci. U.S.A.* **101** (2004), pp. 5573–5578.
- LINDAHL AND NYBERG, 1972 — T. Lindahl and B. Nyberg, Rate of depurination of native DNA, *Biochemistry* **11** (1972), pp. 3610–3618.

- LINDAHL AND NYBERG, 1974 — T. Lindahl and B. Nyberg, Heat-induced deamination of cytosine residues in deoxyribonucleic acid, *Biochemistry* **13** (1974), pp. 3405–3410.
- MEHTA *ET AL.*, 1997 — D.V. Mehta, R.J. DiGate, D.L. Banville and R.D. Guiles, Optimized gene synthesis, high level expression, isotopic enrichment, and refolding of Human Interleukin-5, *Prot. Expr. Purif.* **11** (1997), pp. 86–94.
- MIZUGUCHI *ET AL.*, 1999 — H. Mizuguchi, M. Nakatsuji, S. Fujiwara, M. Takagi and T. Imanaka, Characterization and application to hot start PCR of neutralizing monoclonal antibodies against KOD DNA polymerase, *J. Biochem.* **126** (1999), pp. 762–768.
- MOORE, 2005 — P. Moore, PCR: replicating success, *Nature* **435** (2005), pp. 235–238.
- QUINTANAR AND NELSON, 2002 — Quintanar, A., Nelson, R.M., 2002. A process and apparatus for high-speed amplification of DNA. U.S. Patent No. 6,472,186.
- PADHYE *ET AL.*, 2002 — N.V. Padhye, R.M. Nelson, A. Quintanar, H.J. Viljoen, S.E. Whitney, D. Shoemaker and E.A. Henchal, High-speed amplification of *Bacillus anthracis* DNA using a helium-CO<sub>2</sub> gas thermocycler, *Genet. Eng. News* **22** (11) (2002), pp. 42–43.
- PIENAAR *ET AL.*, 2006 — E. Pienaar, M. Theron, M. Nelson and H.J. Viljoen, A quantitative model of error accumulation during PCR amplification, *Comput. Biol. Chem.* **30** (2006), pp. 102–111.
- ROUILLARD *ET AL.*, 2004 — J.M. Rouillard, W. Lee, G. Truan, X. Gao, X. Zhou and E. Gulari, Gene2Oligo: oligonucleotide design for in vitro gene synthesis, *Nucleic Acids Res.* **32** (2004), pp. W176–W180.
- SAIKI *ET AL.*, 1985 — R.K. Saiki, S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich and N. Anheim, Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia, *Science* **230** (1985), pp. 1350–1354.
- SHEVCHUK *ET AL.*, 2004 — N.A. Shevchuk, A.V. Bryksin, Y.A. Nusinovich, F.C. Cabello, M. Sutherland and S. Ladisch, Construction of long DNA molecules using long PCR-based fusion of several fragments simultaneously, *Nucleic Acids Res.* **32** (2004), pp. e19–e47.
- SIMMONDS AND LANE, 1999 — R.E. Simmonds and D.A. Lane, Structural and functional implications of the intron/exon organization of the human endothelial cell protein C/activated protein C receptor (EPCR) gene: comparison with the structure of CD1/major histocompatibility complex alpha1 and alpha2 domains, *Blood* **94** (1999), pp. 632–641.
- SMITH *ET AL.*, 2003 — H.O. Smith, C.A. Hutchison, C. Pfannkock and J.C. Venter, Generating a synthetic genome by whole genome assembly: ΦX174 bacteriophage from synthetic oligonucleotides, *Proc. Natl. Acad. Sci. U.S.A.* **100** (2003), pp. 15,440–15,445.
- STEMMER *ET AL.*, 1995 — W.P.C. Stemmer, A. Cramer, K.D. Ha, T.M. Brennan and H.L. Heyneker, Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides, *Gene* **164** (1995), pp. 49–53.
- TIAN *ET AL.*, 2004 — J. Tian, H. Gong, N. Sheng, X. Zhou, E. Gulari, X. Gao and G. Church, Accurate multiplex gene synthesis from programmable DNA microchips, *Nature* **432** (2004), pp. 1050–1053.
- WHITNEY *ET AL.*, 2004 — S.E. Whitney, A. Sudhir, R.M. Nelson and H.J. Viljoen, Principles of rapid polymerase chain reactions: mathematical modeling and experimental verification, *Comput. Biol. Chem.* **28** (2004), pp. 195–209.
- WITTWER *ET AL.*, 1990 — C.T. Wittwer, G.C. Filmore and D.J. Garling, Minimizing the time required for DNA amplification by efficient heat transfer to small samples, *Anal. Biochem.* **186** (1990), pp. 328–331.
- YOUNG AND DONG, 2004 — L. Young and Q. Dong, Two-step total gene synthesis method, *Nucleic Acids Res.* **32** (2004), pp. e59–e71.

○ ○ ○ ○

Supplemental data, Table S, follows.

Table S. Oligonucleotides for EPCR gene 1 and 2 as designed by Gene2Oligo software.

---

R0	TCGAGATTCGCCCAGCC
F0	GGCTGGGCGAATCTCGAGAAAAGATTTTGTAGCCAAGACGCC
R17	AAGTCTTTGGAGGCCATCTGAGGCGTCTTGGCTACAAAATCTTTTC
F42	TCAGATGGCCTCCAAAGACTTCATATGCTCCAGATCTCCTACTTCC
R63	ACACGTGATAGGGGTCGCGGAAGTAGGAGATCTGGAGCATATG
F88	GCGACCCCTATCACGTGTGGTACCAGGGCAACGCG
R106	TAGGTGTCCCCCAGCGACGCGTTGCCCTGGTACC
F123	TCGCTGGGGGGACACCTAACGCACGTGCTGGAAGG
R141	CGTGGTGTGGTGTCTGGGCCTTCCAGCACGTGCGT
F158	CCCAGACACCAACACCACGATCATTGAGCTGCAGCCCT
R177	GCTCTCGGGCTCCTGCAAGGGCTGCAGCTGAATGAT
F196	TGCAGGAGCCCGAGAGCTGGGCGCGCACGCAGAG
R213	GGTAGGACTGCAGGCCACTCTGCGTGCGCGCCCA
F230	TGGCCTGCAGTCCTACCTGCTCCAGTTCCACGGC
R247	TGCACCAGGCGCACGAGGCCGTGGAAGTGGAGCA
F264	CTCGTGCGCCTGGTGCACCAGGAGCGGACCTTGG
R281	GCGGATGGTCAGAGGAAAGGCCAAGGTCCGCTCCTGG
F298	CCTTTCCTCTGACCATCCGCTGCTTCCTGGGCTGTGA
R318	AGCCCTCGGGAGGCAGCTCACAGCCCAGGAAGCA
F335	GCTGCCTCCCGAGGGCTCTAGAGCCCATGTCTTCTTCGA
R352	GCTCCCATTCACAGCCACTTCGAAGAAGACATGGGCTCTAG
F374	AGTGGCTGTGAATGGGAGCTCCTTTGTGAGTTTCCGGCC
R393	TGCCACAAGGCTCTCTCCGGCCGGAAACTCACAAAGGA
F413	GGAGAGAGCCTTGTGGCAGGCAGACACCCAGGTCA
R431	GGTGACCACTCCGGAGGTGACCTGGGTGTCTGCC
F448	CCTCCGGAGTGGTCACCTTCACCCTGCAGCAGCT
R465	GAGTGCGGTTGTAGGCATTGAGCTGCTGCAGGGTGAA
F482	CAATGCCTACAACCGCACTCGGTATGAACTGCGGGAATTCCT
R502	TGCACACAGGTGTCCTCCAGGAATTCCTGCAGTTCATACC
F524	GGAGGACACCTGTGTGCAGTATGTGCAGAAACATATTTCCGCG
R542	TTGGCTCCCTTTTCGTGTTTTCCGCGGAAATATGTTTCTGCACATAC
F567	GAAAACACGAAAGGGAGCCAAACAAGCCGCTCCTACACTT
R588	AGATAGCGGCCGCTTACGAAGTGTAGGAGCGGCTTGT
F607	CGTAAGCGGCCGCTATCTGTTGGGCCCTGCTGGCA
F625	TGCCAGCAGGGCCCAAC

Table S. (continued)

---

R0	AATCTTTTCTCGAGATGTGCTCGG
F0	CCGAGCACATCTCGAGAAAAGATTTTGTAGCCAAGACGCCTCA
R24	CATATGAAGTCTTTGGAGGCCATCTGAGGCGTCTTGGCTACAA
F43	GATGGCCTCCAAAGACTTCATATGCTCCAGATCTCCTACTTCCGC
R67	ACCACACGTGATAGGGGTCGCGGAAGTAGGAGATCTGGAG
F88	GACCCCTATCACGTGTGGTACCAGGGCCTGGGGGGA
R107	CAGCACGTGCGTTAGGTGTCCCCCAGGCCCTGGT
F124	CACCTAACGCACGTGCTGGAAGGCCCAGACACCATCA
R142	GCAAGGGCTGCAGCTGAATGATGGTGTCTGGGCCTTC
F161	TTCAGCTGCAGCCCTTGCAGGAGCCCGAGAGCTGG
R179	CCACTCTGCGTGCGCGCCAGCTCTCGGGCTCCT
F196	GCGCGCACGCAGAGTGGCCTGCAGTCCTACCTGCT
R213	CGAGGCCGTGGAAGTGGAGCAGGTAGGACTGCAGG
F231	CCAGTTCCACGGCCTCGTGCGCCTGGTGCACCAG
R248	AAGGCCAAGGTCCGCTCCTGGTGCACCAGGCGCA
F265	GAGCGGACCTTGGCCTTTCCTCTGACCATCCGCTG
R282	AGCTCACAGCCCAGGAAGCAGCGGATGGTCAGAGGA
F300	CTTCCTGGGCTGTGAGCTGCCTCCCGAGGGCTCTA
R318	ACTTCGAAGAAGACATGGGCTCTAGAGCCCTCGGGAGGC
F335	GAGCCCATGTCTTCTTCGAAGTGGCTGTGTCCTTTGTGAGTTT
R357	AGGCTCTCTCCGGCCGGAAACTCACAAAGGACACAGCC
F378	CCGGCCGGAGAGAGCCTTGTGGCAGGCAGACACC
R395	CACTCCGGAGGTGACCTGGGTGTCTGCCTGCCACA
F412	CAGGTCACCTCCGGAGTGGTCACCTTCACCCTGCAGC
R430	ATACCGGTAGGCATTGAGCTGCTGCAGGGTGAAGGTGAC
F449	AGCTCAATGCCTACCGGTATGAACTGCGGGAATTCCTGGA
R469	ACATACTGCACACAGGTGTCCTCCAGGAATTCCCGCAGTTC
F489	GGACACCTGTGTGCAGTATGTGCAGAAACATATTTCCGCGGA
R510	GTTTGGCTCCCTTTCGTGTTTTCCGCGGAAATATGTTTCTGC
F531	AAACACGAAAGGGAGCCAAACAAGCCGCTCCTACACTTCG
R552	TCGCTAGCGGCCGCTTACGAAGTGTAGGAGCGGCTT
F571	TAAGCGGCCGCTAGCGAGGAGTAGCTGCTTCCCGT
F588	ACGGGAAGCAGCTACTCC