

Experimental analysis of gene assembly with TopDown one-step real-time gene synthesis

Hongye Ye, Mo Chao Huang, Mo-Huang Li* and Jackie Y. Ying

Institute of Bioengineering and Nanotechnology, 31 Biopolis Way, The Nanos, Singapore 138669

Received January 8, 2009; Accepted February 11, 2009

ABSTRACT

Herein we present a simple, cost-effective TopDown (TD) gene synthesis method that eliminates the interference between the polymerase chain reactions (PCR) assembly and amplification in one-step gene synthesis. The method involves two key steps: (i) design of outer primers and assembly oligonucleotide set with a melting temperature difference of $>10^{\circ}\text{C}$ and (ii) utilization of annealing temperatures to selectively control the efficiencies of oligonucleotide assembly and full-length template amplification. In addition, we have combined the proposed method with real-time PCR to analyze the step-wise efficiency and the kinetics of the gene synthesis process. Gel electrophoresis results are compared with real-time fluorescence signals to investigate the effects of oligonucleotide concentration, outer primer concentration, stringency of annealing temperature, and number of PCR cycles. Analysis of the experimental results has led to insights into the gene synthesis process. We further discuss the conditions for preventing the formation of spurious DNA products. The TD real-time gene synthesis method provides a simple and efficient method for assembling fairly long DNA sequence, and aids in optimizing gene synthesis conditions. To our knowledge, this is the first report that utilizes real-time PCR for gene synthesis.

INTRODUCTION

Synthetic gene with man-made DNA sequences has become a powerful molecular tool with broad applications on engineering proteins (1,2), artificial gene networks (3) and synthetic genomes (4–6). The DNA biomolecules are constructed by assembling pools of oligonucleotides into larger DNA using *de novo* PCR- (6,7) or ligase chain reaction (LCR)-based (4,8) synthesis methods. The most reported methods for constructing long DNA were based on the PCR process, which relied on the use of

overlapped oligonucleotides to construct genes. Various PCR-based methods have been proposed in attempt to optimize the PCR process for long DNA sequences, and to enhance the accuracy of assembly. These methods include the thermodynamically balanced inside-out (TBIO) method (9), successive PCR (10), dual asymmetrical PCR (DA-PCR) (11), overlap extension PCR (OE-PCR) (12,13), PCR-based two-step DNA synthesis (10,14,15) and one-step gene synthesis (16). While various PCR approaches have been implemented, mysteriously the reported optimal synthesis conditions always coincided with an oligonucleotide concentration of 10–60 nM, an outer primer concentration of 200–800 nM, and a PCR cycle number of 20–35 (9,10,15–18). Genes assembled from unpurified, desalted oligonucleotides often resulted in full-length sequence with spurious products of higher molecular weights (9–12,16–19), which degraded the purity of the synthesized products. This abnormal event is usually neglected without explanation by authors. Although PCR-based gene synthesis has been widely implemented, there is a lack of capability in accurately predicting the gene synthesis. The development of an accurate model of gene synthesis would aid in understanding these phenomena and designing optimal reaction conditions.

Another issue associated with gene synthesis is the lack of a standard or universal method (20). Depending on the complexity of target genes, the synthetic genes are often constructed with a one-step or two-step overlapping process. For its simplicity, the one-step process is preferred for short DNAs (<500 bp), wherein the amplification primers are mixed with assembly oligonucleotides in a single PCR reaction. The melting temperatures (T_m) of these oligonucleotides are commonly designed with homologous T_m to balance the PCR assembly and amplification. Hence, the excess outer primers would preferentially anneal with extended oligonucleotides, resulting in a potentially large portion of amplification primers participating in the initial gene assembly cycles. This could deplete the deoxynucleotide triphosphates (dNTPs) and cease the PCR reaction (17,21). This competitive effect is more critical for DNA with high GC content or length (9,10), and is eliminated in the two-step PCR process

*To whom correspondence should be addressed. Tel: +65 6824 7119; Fax: +65 6478 9085; Email: mhli@ibn.a-star.edu.sg

whereby the amplification and assembly are performed separately with the extra cost of fresh PCR mixture and additional human effort. In the two-step process, a pool of short oligonucleotides is first assembled into long double-stranded DNA (dsDNA) (termed 'template') with the desired length and sequence information using the polymerase chain assembly (PCA). The quantity of the assembled template DNA is then amplified by the PCR step.

Herein we described a novel approach that combined the simplicity and cost effectiveness of the one-step process, with the assembly efficiency of the two-step process for the synthesis of relatively long genes. This method utilized software to design primers and oligonucleotides with distinct melting temperatures to minimize the competition between the PCA assembly and PCR amplification in the one-step gene synthesis. Figure 1 shows the concept of the proposed TopDown (TD) one-step gene assembly method. The outer primers ($T_m \sim 50^\circ\text{C}$) and inner oligonucleotides ($T_m \sim 65^\circ\text{C}$) were designed with a melting temperature difference (ΔT_m) of $\sim 15^\circ\text{C}$. The overlapping gene synthesis was conducted in one PCR mixture with two annealing temperature segments for oligonucleotide assembly and full-length template amplification, respectively. A high annealing temperature (T_{ah}) (matched to the T_m of oligonucleotides) was used for the first 20 cycles aimed at assembly. This was followed by a low annealing temperature (matched to the T_m of primers) to boost the full-length amplification. The outer primers were subjected to an elevated annealing condition ($T_{ah} - T_m = 15^\circ\text{C}$) during the early assembly process, which prevented mis-pairing among primers and oligonucleotides.

We further combined the TD one-step gene synthesis with real-time PCR to investigate the gene synthesis process. LCGreen I and SYBR Green I intercalation dyes were examined for real-time gene synthesis. Gel electrophoresis results were compared with the real-time fluorescence signals to study the effects of the oligonucleotide concentration, outer primer concentration, stringency of annealing temperature, and number of PCR cycles. Performing real-time gene synthesis provided a novel approach to explore these factors in a systematic study, which revealed unique information not available from gel electrophoresis results (16). An analytical model was further developed, which uncovered the mystery of coincident optimal conditions, and the formation of undesired, spurious DNA products. This real-time method provided new insights into gene synthesis and a universal method for gene synthesis.

MATERIALS AND METHODS

Design of oligonucleotides for gene synthesis

Gene sequence for the promoter of human calcium-binding protein A4 (S100A4, 752 bp; chr1:1503312036-1503311284) (22) was selected for synthesis via assembly PCR. Oligonucleotides were derived by a custom-developed program, which first divided the given sequence into oligonucleotides of approximately equal lengths by

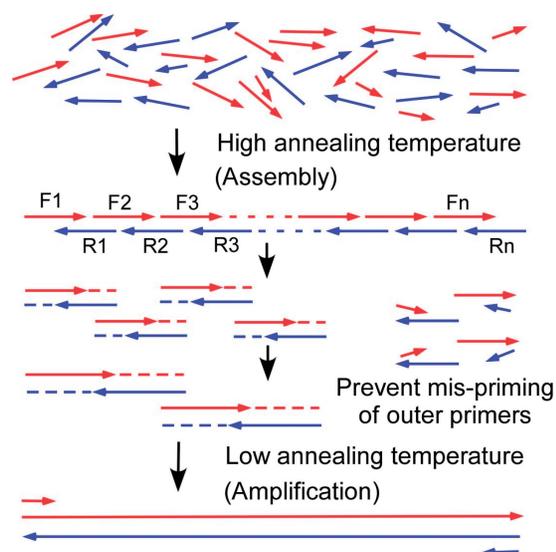


Figure 1. Schematic illustration of TD one-step gene synthesis combining PCR assembly and amplification into a single stage with different annealing temperatures designed for assembly and amplification. Inner oligonucleotides and outer primers are designed with a melting temperature difference of $>15^\circ\text{C}$ to minimize potential interference during PCR.

markers, and computed the average and deviation in melting temperatures among the overlapping regions using the nearest-neighbor model with SantaLucia's thermodynamic parameter (23), corrected with salt and oligonucleotide concentrations. Next, the oligonucleotide lengths were adjusted through shifting the marker positions to minimize the deviations in the overall overlapping melting temperature. The summary of oligonucleotide set is shown in Table 1 with the detail information provided in Supplementary Table S1.

TopDown one-step real-time gene synthesis

TopDown one-step process was optimized using real-time PCR conducted with Roche's LightCycler 1.5 real-time thermal cycling machine with a temperature transition of $20^\circ\text{C}/\text{s}$. Real-time gene synthesis was conducted with $20\ \mu\text{l}$ of reaction mixture including $1\times$ PCR buffer (Novagen), $1\ \mu\text{l}$ of $0.25\times$ to $5\times$ SYBR Green I ($1\times = 1/20\ 000$ dilution; Invitrogen) or LCGreen I (Idaho Technology Inc.), $4\ \text{mM}$ of MgSO_4 , $1\ \text{mM}$ each of dNTP (Stratagene), $500\ \mu\text{g}/\text{ml}$ of bovine serum albumin (BSA), $5\text{--}80\ \text{nM}$ of oligonucleotides, $60\ \text{nM}$ to $1\ \mu\text{M}$ of forward and reverse primers and $1\ \text{U}$ of KOD Hot Start (Novagen). The PCR were conducted under the following conditions: 2 min of initial denaturation at 95°C ; 20 cycles of 95°C for 5 s, $58\text{--}70^\circ\text{C}$ for 10 s, 72°C for 30 s; followed by 20 cycles of 95°C for 5 s, 49°C for 10 s, 72°C for 30 s; and final extension at 72°C for 10 min. Desalted oligonucleotides were purchased from Research Biolabs (Singapore) and Proligo (Singapore) without additional purification.

One-step and two-step PCR-based gene synthesis

Conventional gene synthesis via PCR was performed either as a one-step process, combining PCR assembly

Table 1. Data of oligonucleotide set

Gene	Length (bp)	Average T_m (min, max) ($^{\circ}\text{C}$)	Number of oligonucleotides	Overlap length (nt)	Oligonucleotide length (nt)
S100A4	752	66.0 (61.1, 69.8)	30	19–33	19, 41–66

Table 2. PCR conditions for one-step, two-step and TD one-step gene syntheses

Method	PCA ^a		PCR	
	Annealing temperature	Number of cycles	Annealing temperature	Number of cycles
One-step	58 $^{\circ}\text{C}$ ^b	30	–	–
Two-step	58 $^{\circ}\text{C}$	30	49 $^{\circ}\text{C}$ ^c	30
TD one-step	58–70 $^{\circ}\text{C}$	20	49 $^{\circ}\text{C}$ ^c	20

^aAverage melting temperature of oligonucleotide set: 66.0 $^{\circ}\text{C}$.

^bMelting temperature of outer primers: 59.4 $^{\circ}\text{C}$ and 63.4 $^{\circ}\text{C}$.

^cMelting temperature of outer primers: 49.4 $^{\circ}\text{C}$ and 50.9 $^{\circ}\text{C}$.

and amplification into a single stage, or as a two-step process with separate stages for assembly and amplification. All PCR reactions, whether for assembly or amplification, were run in standard 0.2 ml PCR tubes with a commercial thermal cycler (DNA Engine PTC-200, Bio-Rad) using the same oligonucleotides set and out primers as in the non-competitive one-step PCR. The one-step process was performed with 50 μl of reaction mixture including 1 \times PCR buffer (Novagen), 4 mM of MgSO_4 , 1 mM each of dNTP (Stratagene), 500 $\mu\text{g}/\text{ml}$ of BSA, 10 nM of oligonucleotides, 400 nM of forward and reverse primers and 1 U of KOD Hot Start (Novagen). The one-step PCR was conducted under the following conditions: 2 min initial denaturation at 95 $^{\circ}\text{C}$; 30 cycles of 95 $^{\circ}\text{C}$ for 5 s, 58 $^{\circ}\text{C}$ for 10 s, 72 $^{\circ}\text{C}$ for 30 s; and final extension at 72 $^{\circ}\text{C}$ for 10 min. The PCR protocol of the two-step process was essentially the same as that for one-step process except for the concentration of oligonucleotides and annealing temperature. For PCR assembly, 10 nM of oligonucleotides were used without the forward and reverse primers. For gene amplification, 2 μl of the assembled product was diluted in 25 μl of amplification reaction mixture with primers concentration of 400 nM each, and an annealing temperature of 49 $^{\circ}\text{C}$ was employed. The PCR conditions of these three gene syntheses are summarized in Table 2.

Agarose gel electrophoresis

The synthesized products were analyzed by 1.5% agarose gel (NuSieve[®] GTG[®], Cambrex Corporation), stained with ethidium bromide (Bio-Rad Laboratories) or SYBR Green (Invitrogen), and visualized using Typhoon 9410 variable imager (Amersham Biosciences). Gel electrophoreses were performed at 100 V for 45 min with 100 bp ladder (New England) and 5 μl of DNA samples.

RESULTS

Performance of TD one-step gene synthesis

Successful gene synthesis was achieved using TD one-step and two-step processes, while no obvious full-length gene product was obtained in one-step PCR process as shown by gel electrophoresis (Figure 2). The TD one-step process was conducted with an annealing temperature (T_{ah}) of 67 $^{\circ}\text{C}$ (average T_m of oligonucleotides = 66 $^{\circ}\text{C}$) for the first 20 cycles, followed by an annealing temperature of 49 $^{\circ}\text{C}$ (average T_m of primers = 50.1 $^{\circ}\text{C}$) for another 20 cycles. The continuous fluorescence monitoring revealed the efficiency of the gene synthesis process (Figure 3). Unlike the exponential nature of PCR amplification, the assembly efficiency was more likely linear in nature.

Two intercalating fluorescent dyes (SYBR Green I and LCGreen I) were investigated for real-time gene synthesis (Supplementary Figure S1). The LCGreen I (24) was more suitable for studying the real-time gene synthesis, which has a fluorescence spectrum similar to the commonly adopted SYBR Green I in real-time PCR. The SYBR Green I would bind preferentially to long DNA fragments (25,26), and could redistribute from short DNA fragments to long DNA fragments during thermal cycling (27). This would make it difficult to analyze the fluorescence signal since the PCA mixture would contain various lengths of dsDNA.

The initial DNA quantity (~6 pmol; 10 nM \times 20 μl \times 30 oligonucleotides) in PCA mixture was much larger (by >6 orders of magnitude) than that in standard PCR amplification (<10⁶ copies of template DNA) (28). The real-time PCR conditions were adjusted for this factor. The optimal concentration of LCGreen I was studied and increased from 1 \times for standard PCR to 2 \times (Supplementary Figure S1). The dNTPs concentration was adjusted from 0.2 mM each for standard PCR to 1 mM each to prevent the depletion of dNTPs. The Mg^{2+} ion (MgSO_4) concentration has been empirically optimized (at 4 mM) based on the concentration of dNTP, which could chelate with Mg^{2+} and affect the polymerase activity (29,30) (Supplementary Figure S2). The manufacturer's recommended Mg^{2+} ion concentration was 1.5 mM for standard PCR with 0.2 mM of dNTPs each.

Analysis of real-time gene synthesis

Mechanistically, gene synthesis took place in several phases, as revealed by the variation in slopes with the number of PCR cycles (Figure 4). This phenomenon was remarkable with an oligonucleotide concentration of 10–20 nM. In the early cycles of PCA, most annealing between paired oligonucleotides formed an extendable duplex, which could undergo extension by polymerase

(phase 1; cycles less than 7). The fluorescence signal revealed a linear increment of DNA length extension with each cycle. In contrast to that reported by Wu *et al.* (16) and Lee *et al.* (21), the assembly efficiency increased with further PCR cycles (phase 2; cycles ~7–14). Our hypothesis was that the assembly process switched in favor of full-length template amplification as the full-length fragments emerged, and was promoted by the excess outer primers. The PCA reaction then reached the first plateau (phase 3; cycles 15–20) whereby the outer primers priming was limited by the elevated annealing conditions ($T_{ah} - T_m = 15^\circ\text{C}$). At cycle 21, the annealing temperature was reduced to 49°C to match with the T_m of primers (phase 4; cycles ~21–29). The exponential amplification was boosted, and caused a sudden jump in fluorescence signals. Finally, the process reached the second plateau, presumably due to the depletion of outer primers or non-specific products annealing (phase 5). The plateau stages were delayed or completely missing for low oligonucleotide concentration ($<7\text{ nM}$) due to its low assembly efficiency.

For gene synthesis with $>64\text{ nM}$ of oligonucleotides, the PCR process reached the plateau within 15 cycles. Additional cycles would most likely favor non-specific PCR, and lead to the generation of spurious bands and the buildup of high molecular weight products in gel electrophoresis (Figure 4b), as observed in most reported gene synthesis results (9–12,16–19). The consistent gel results and real-time PCR curves suggested that the optimal oligonucleotide concentration was 10–20 nM for TD gene synthesis, which coincided with that of both the one-step (16,17) and two-step (18) processes.

We further investigated the effect of primer by varying primer concentration from 60 nM to $1\ \mu\text{M}$ while keeping the oligonucleotide concentration at 10 nM. The highest full-length quantity was obtained with 400 nM of primers

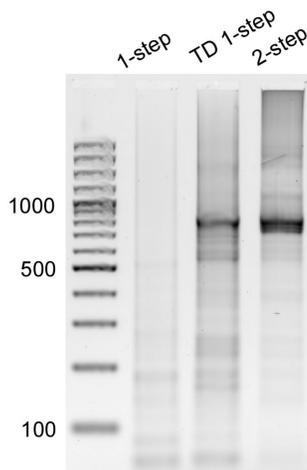


Figure 2. Agarose gel electrophoresis results of one-step (30 cycles), TD one-step (40 cycles) and two-step (PCA: 30 cycles; PCR: 30 cycles) gene syntheses. The TD one-step process is conducted with 20 cycles with an annealing temperature of 67°C , followed by 20 cycles with an annealing temperature of 49°C . The concentrations of oligonucleotides and outer primers are 10 nM and 400 nM, respectively.

(Figure 5), which was consistent with observations in one-step (16) and two-step (18) processes. Assembly efficiencies, depicted by the slopes of fluorescence increment, were indifferent in the early cycles (less than cycle 7),

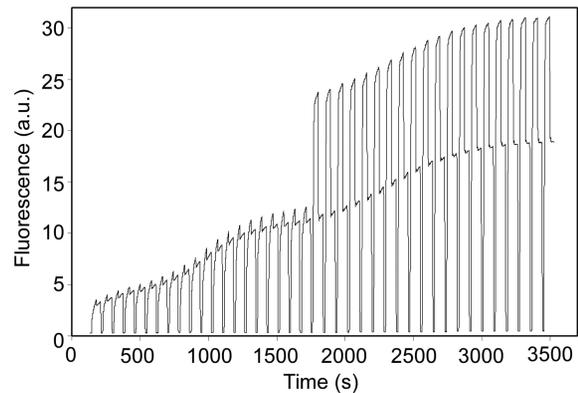


Figure 3. Continuous fluorescence monitoring of real-time gene synthesis with $1\times$ LCGreen I. The first 20 cycles are conducted with an annealing temperature of 67°C , and the next 20 cycles are conducted with an annealing temperature of 49°C . The concentrations of oligonucleotides and outer primers are 10 nM and 400 nM, respectively.

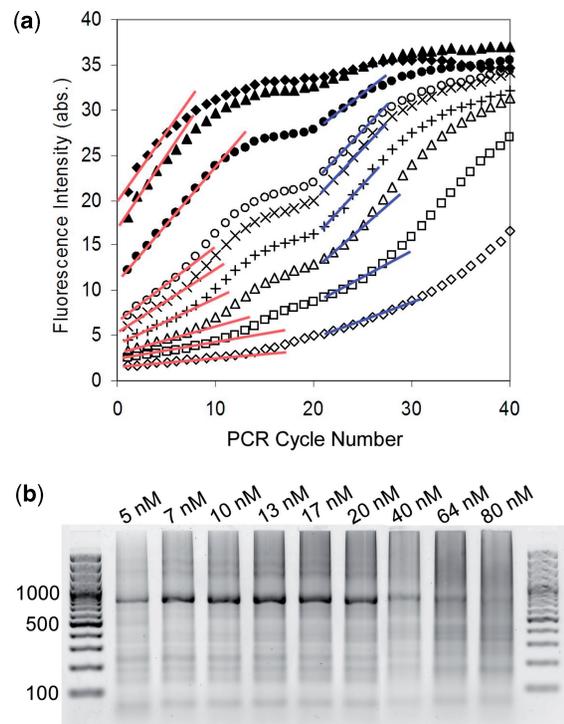


Figure 4. The oligonucleotide concentration is critical in the successful gene synthesis. S100A4 (752 bp) is synthesized with various oligonucleotide concentrations ranging from 5 nM to 80 nM, and annealing temperatures of 67°C for the first 20 cycles and 49°C for the next 20 cycles. (a) Fluorescence as a function of PCR cycle number for oligonucleotide concentrations of 5 nM (unfilled diamond), 7 nM (unfilled square), 10 nM (unfilled triangle), 13 nM (+), 17 nM (\times), 20 nM (unfilled circle), 40 nM (filled circle), 64 nM (filled triangle) and 80 nM (filled diamond). The slopes of fluorescence increment in the early cycles and cycles #21 indicate the efficiencies of the assembly and amplification processes. (b) The corresponding agarose gel electrophoresis results.

even though the primer concentration was varied by 16-fold (inset in Figure 5a). This demonstrated the non-interference feature of the TD process, wherein the outer primers did not intervene with the assembly process. The assembly efficiencies started to deviate at around cycle 8 as the full-length products emerged, in favor of full-length template amplification. Unlike the oligonucleotide concentration, which dominated the assembly reaction and critically influenced the success of gene synthesis, the primer concentration was less critical. It presumably controlled the late amplification process and the quantity of desired DNA. The optimal PCR cycles depended on the initial oligonucleotide concentration and target gene length. This was clearly demonstrated by the experiment on oligonucleotide concentration (Figure 4). As oligonucleotide concentration increased from 20 nM to

80 nM, the full-length band gradually disappeared and became widened.

The overlapping assembly was a parallel process. Relatively few PCR cycles were needed to complete the assembly. The theoretical minimum number of cycles (x) in order to construct a dsDNA molecule of length (L) from uniform oligonucleotide length (n) and overlapping size (s) is given by:

$$2^x n - (2^x - 1) s > L$$

Theoretically, six PCA cycles were sufficient for assembling S100A4 (752 bp) from a pool of 40 nt oligonucleotides with an overlap of 20 nt. To determine whether excess cycling was necessary for gene assembly, we used the optimal condition determined in previous experiments with various PCA cycles of 6–20, followed by 20 amplification cycles. Gene synthesis was fairly efficient. Indeed, full-length assembly was achieved within 11 PCA cycles (Figure 6).

The gene synthesis was insensitive to the variation in assembly annealing temperature (T_{ah}) from 58°C to 70°C, as visualized in both gel results and fluorescence signals (Figure 7). The fluorescence intensity curves were indiscriminant to the annealing temperatures during the assembly phase (first 13 cycles), and began to deviate only after the first phase (see inset in Figure 7a). The indifference in fluorescence intensity during the first 13 cycles implied that the primers did not intervene with the assembly reaction. The primers were designed with an average T_m of 50.9°C, which meant that the primers encountered an annealing stringent of 7.1–19.9°C ($T_{ah} - T_m$) during the PCA process. This suggested that we could potentially reduce the melting temperature window (ΔT_m of primers and oligonucleotides) to 7.1°C, and ensure the non-competitive feature of TD gene synthesis method. Interestingly, a higher yield of the desired DNA was obtained with a stringent annealing temperature (>67°C) higher than the average T_m of oligonucleotides (66°C).

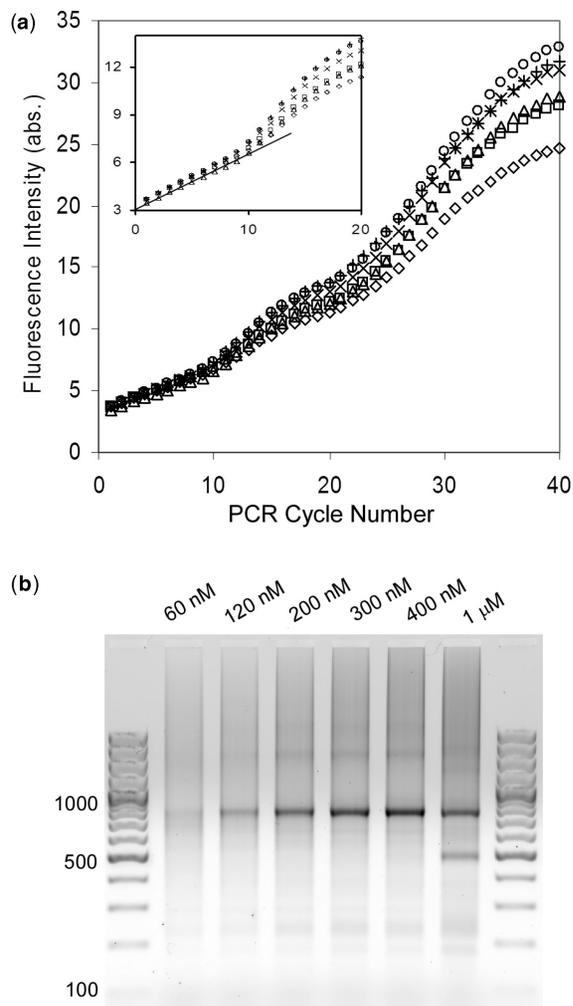


Figure 5. S100A4 (752 bp) is successfully synthesized with various primer concentrations ranging from 60 nM to 1 μM, as indicated by the sharp, narrow gel band of the desired length. (a) Fluorescence as a function of PCR cycle number for outer primer concentrations of 60 nM (unfilled diamond), 120 nM (unfilled square), 200 nM (unfilled triangle), 300 nM (x), 400 nM (+) and 1 μM (unfilled circle). The inset shows the fluorescence signal of the first 20 cycles. (b) The corresponding agarose gel electrophoresis results.

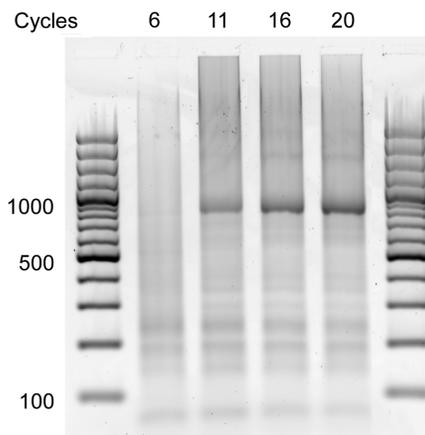


Figure 6. S100A4 is synthesized with various assembly cycles (6–20 cycles), followed by another 20 cycles for amplification. Agarose gel electrophoresis results indicate that full-length assembly is achieved within 11 cycles.

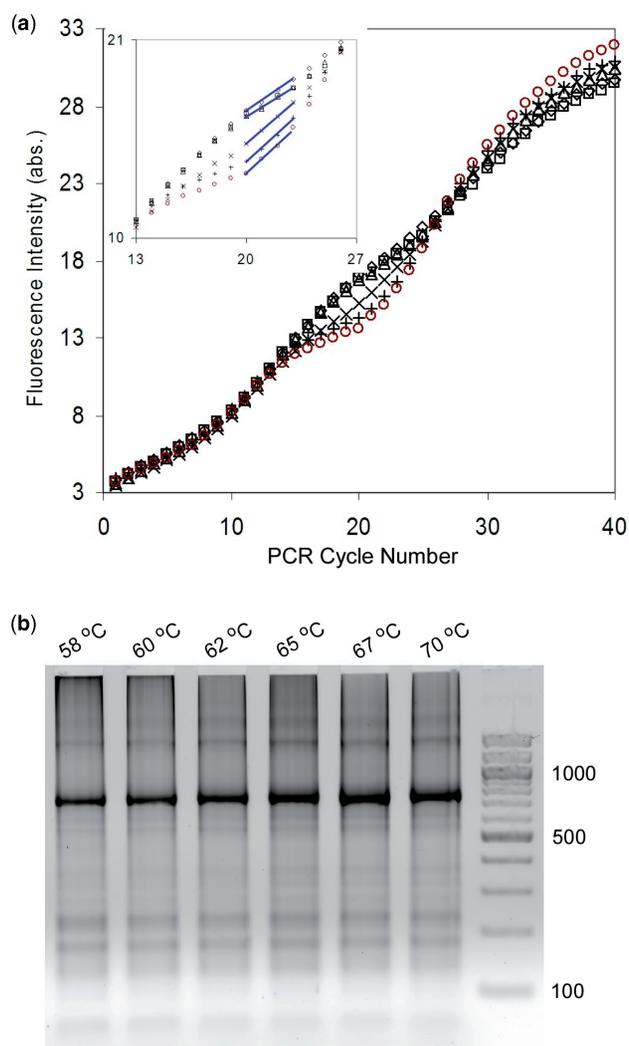


Figure 7. S100A4 (752 bp) synthesized with various assembly annealing temperatures ranging from 58°C to 70°C for the first 20 cycles, followed by an annealing temperature of 49°C for the next 20 cycles. (a) Fluorescence as a function of PCR cycle number for annealing temperatures of 58°C (unfilled diamond), 60°C (unfilled square), 62°C (unfilled triangle), 65°C (x), 67°C (+) and 70°C (unfilled circle). The inset shows the middle 15 cycles (no. 13–27). (b) The corresponding agarose gel electrophoresis results. Higher synthesis yield is obtained with a stringent assembly annealing temperature (>67°C).

DISCUSSION

We have presented the TD gene synthesis method, which offered a simple, rapid and low-cost method for synthesizing fairly long DNA (752 bp) with only one PCR step. The primer interference problem present in one-step process has been eliminated by designing primers and assembly oligonucleotide set with melting temperature variation (ΔT_m) of >15°C, and performing T_m -matched PCR to selectively control the efficiencies of oligonucleotide assembly and full-length template amplification. Our data suggested that the TD process might also work well with a relaxed ΔT_m of 7.1°C. The TD synthesis conditions have been empirically optimized using real-time PCR with LCGreen I dye. It was noted that the popular

SYBR Green I dye was not suitable for real-time gene synthesis. Other intercalating dyes such as LCGreen (25) and SYTO9 (31) might also work, but required further optimization in dye concentration.

Based on the data presented herein, we have developed an analytical model that described the TD one-step gene synthesis process clearly. The TD synthesis occurred in five phases in term of PCR cycles: (i) linear assembly, (ii) emerging amplification, (iii) first plateau, (iv) exponential amplification and (v) non-specific amplification. Our experiments have demonstrated that TD one-step gene synthesis was fairly efficient, as compared to the conventional one-step and two-step processes. We found the assembly process automatically switched to preferential full-length amplification as the full-length template emerged. This greatly improved the assembly efficiency of the PCA process as compared to the conventional one-step and two-step processes.

We found that the quality and quantity of PCR-based gene synthesis were influenced by several factors, including annealing temperature, concentration of oligonucleotides, concentration of monomers and number of PCR cycles. The fluorescence curve (10 nM curve in Figure 4) suggested the assembly and amplification processes reached the plateau at around cycle 15 and cycle 35, respectively. This implied that the optimal PCR cycles were 30 cycles: 15 cycles each for assembly and amplification reactions. Furthermore, after the PCR reached the plateau, the amplification efficiency of PCR decreased. Additional PCR cycling would favor the non-specific annealing of the full-length product to assembled random-fragments or the full-length product itself. Both extendable and unextendable pairings could occur. Products annealed in the 3' recessed configuration could be extended to higher molecular weight DNA and be randomly terminated during the later cycles. In contrast, products annealed with 3' ends protruded would reduce the amount of full-length DNA, and generate DNA with lower molecular weight. Both of these types of non-specific annealing would result in diminished full-length gel band (32), as observed in gel electrophoresis. These abnormal products with incorrect DNA sequences would potentially complicate the enzymatic cleavage or consensus shuffling error correction process (33,34) when consecutive PCRs were adopted for gene re-assembly or DNA re-amplification.

Although the gel results (Figure 1) indicated that the two-step process was superior to the TD one-step process, we could argue that the optimized TD approach could eventually outperform the two-step process in providing better purity of synthetic products, with the understanding of gene synthesis mechanism. The two-step method suffered the same rapid decrease on extension efficiency as the one-step process during the assembly process (PCA), which created a lot of intermediate DNAs with lower molecular weight besides the full-length products. The non-specific annealing and extension would most likely also occur during PCA, considering the observed fairly efficient full-length assembly (<10 cycles) (Figure 5). These higher molecular weight products could also contain primers' binding sites, and be exponentially

Table 3. Some reported optimal gene synthesis conditions

Name of synthesized DNA	Length (bp)	Oligos (nM)	Primer (nM)	Number of PCR cycles	Method
PDK1 gene	1712	20	200	25–35/25–35	Two-step PCR (9)
PDK1 gene	1712	40–200	200	25/25	TBIO (9)
Various DNAs	470–1200	40	200	20/15	Two-step PCR (12)
<i>vip3aI</i> gene	2370	30	600	25/25	Two-step PCR (10)
Various DNAs	209–936	10–25	400–800	25	One-step PCR (16)
<i>Pur</i> operon	12000	60	600	25/25	PCR-based (15)
Various DNAs	327–993	10–25	500	35–45	One-step PCR (17)
GFP segment	752	5–15	400	30/30	Two-step PCR (18)

amplified at the next amplification step (PCR) as the full-length products (32,35). These incorrect products would complicate the error filtering process (33,36), requiring excess MutS enzyme. We could improve the TD process by having the PCR amplification tailored after the emergence of full-length DNA to avoid excess PCR.

With the help of our model, insights into the coincidental optimal conditions reported for various PCR-based gene synthesis processes were attained (Table 3). The optimal conditions were universal with an oligonucleotide concentration of 10–60 nM and a primer concentration of 200–800 nM, as reported by a distinct gel band of full-length products. We have demonstrated that the success of gene synthesis relied on the assembly efficiency of PCA, which was in turn dominated by the oligonucleotide concentration and number of PCR cycles (Figures 4 and 6). The fluorescence signals indicated that an oligonucleotide concentration of 10–60 nM provided optimal assembly efficiency with full-length products. It was noteworthy that the optimal number of PCR cycles should be adjusted according to the oligonucleotide concentration and gene length to avoid excessive PCR. Successful gene synthesis (752 bp) was also achieved with an oligonucleotide concentration of <5 nM in 40 PCR cycles. However, it might not be reliable for genes with a high GC content or a high length due to the comparatively low assembly efficiency. The primer concentration was less critical; it controlled the late amplification process and the quantity of desired DNA (Figure 5). The optimal primer concentration balanced the quantity and purity of the desired DNA, avoiding the formation of spurious DNA products.

As the synthesis yield was insensitive to the assembly annealing temperature (Figure 7), we suggested performing the assembly with an annealing temperature slightly higher than the average T_m of oligonucleotides. This provided several advantages in (i) eliminating potential competition between the assembly and amplification reactions, (ii) minimizing the possibility of truncated oligonucleotides ($n - 1$) participating in the assembly process and the resulting errors, (iii) providing an stringent annealing condition to reduce the potential of forming secondary structures and (iv) increasing the specialization of oligonucleotides hybridization as in Touchdown PCR (37). All of these would prevent the generation of faulty sequence, especially for gene with high GC contents.

Besides providing a tool for optimizing the gene synthesis conditions, the real-time method would potentially provide other functions for characterizing the synthesis products via analysis of DNA-melting curve. Presumably, successful gene synthesis would yield a product with a single, sharp melting peak, while incomplete synthesis would result in a broad melting curve. The integrated area of the melting peak in the negative derivative of the fluorescence with respect to temperature ($-dF/dT$ versus T) would give the quantity of the desired full-length product (38). In combination with DNA-melting simulation software (31,39), the purity and quantity of assembled products could be estimated with the use of melting curve analysis, which would eliminate the necessity for agarose gel electrophoresis. This approach could be integrated with microfluidic gene synthesis (17,18,40) to develop automated gene synthesis.

Another important factor for successful gene synthesis was the polymerase enzyme. The performance of various polymerases has been studied and compared for the gene synthesis process (2,16,41–43). The KOD series of polymerases were suggested for gene synthesis (2,16,41). We have observed that the KOD Hot Start outperformed the *Taq* and *Pfu* polymerases (data not shown). No obvious full-length gene product was obtained with *Taq* or *Pfu* for the TD method.

Herein we have presented a simple, efficient and cost-effective TD one-step gene synthesis method that combines the advantages of one-step and two-step gene synthesis processes. We recommend conducting TD synthesis with the following conditions: (i) design primers ($T_m = 50–55^\circ\text{C}$) and inner oligonucleotide ($T_m \sim 65^\circ\text{C}$) with distinct melting temperature ($\Delta T_m > 8^\circ\text{C}$); (ii) $2 \times$ LCGreen I, 4 mM of MgSO_4 , 1 mM each of dNTP, 500 $\mu\text{g/ml}$ of BSA, 10 nM of oligonucleotides, 400 nM of forward and reverse primers and 1 U of KOD Hot Start, (iii) 2 min of initial denaturation at 95°C ; 15 cycles of 95°C for 5 s, $67–70^\circ\text{C}$ (according to the T_m of inner oligonucleotides) for 30 s and 72°C for 30 s; followed by 15 cycles of 95°C for 5 s, $50–55^\circ\text{C}$ (according to the T_m of outer primers) for 30 s and 72°C for 30 s.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors acknowledge Drs Shu Wang and Jérôme Boulaire for their helpful suggestions on S100A4.

FUNDING

Institute of Bioengineering and Nanotechnology (Biomedical Research Council, Agency for Science, Technology and Research, Singapore). Funding for open access charge: Institute of Bioengineering and Nanotechnology.

Conflict of interest statement. None declared.

REFERENCES

- Ramachandran,N., Hainsworth,E., Bhullar,B., Eisenstein,S., Rosen,B., Lau,A.Y., Walter,J.C. and Labaer,J. (2004) Self-assembling protein microarrays. *Science*, **305**, 86–90.
- Cox,J.C., Lape,J., Sayed,M.A. and Hellinga,H.W. (2007) Protein fabrication automation. *Protein Sci.*, **16**, 379–390.
- Sprinzak,D. and Elowitz,M.B. (2005) Reconstruction of genetic circuits. *Nature*, **438**, 443–448.
- Smith,H.O., Hutchison,C.A. III, Pfannkoch,C. and Venter,J.C. (2003) Generating a synthetic genome by whole genome assembly: Φ X174 bacteriophage from synthetic oligonucleotides. *Proc. Natl Acad. Sci. USA*, **100**, 15440–15445.
- Gibson,D.G., Benders,G.A., Andrews-Pfannkoch,C., Denisova,E.A., Baden-Tillson,H., Zaveri,J., Stockwell,T.B., Brownley,A., Thomas,D.W., Algire,M.A. *et al.* (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science*, **319**, 1215–1220.
- Cello,J., Paul,A.V. and Wimmer,E. (2002) Chemical synthesis of poliovirus cDNA: Generation of infectious virus in the absence of natural template. *Science*, **297**, 1016–1018.
- Kodumal,S.J., Patel,K.G., Reid,R., Menzella,H.G., Welch,M. and Santi,D.V. (2004) Total synthesis of long DNA sequences: Synthesis of a contiguous 32-kb polyketide synthase gene cluster. *Proc. Natl Acad. Sci. USA*, **101**, 15573–15578.
- Gao,X., Yang,F.Y., Yang,W.J., Lo,S.H. and Kao,C.F. (1998) Gene synthesis by a LCR-based approach: High-level production of leptin-L54 using synthetic gene in *Escherichia coli*. *Biochem. Biophys. Res. Commun.*, **248**, 200–203.
- Gao,X., Yo,P., Keith,A., Ragan,T.J. and Harris,T.K. (2003) 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**, e143.
- Xiong,A.-S., Yao,Q.-H., Peng,R.-H., Li,X., Fan,H.-Q., Cheng,Z.-M. and Li,Y. (2004) A simple, rapid, high-fidelity and cost-effective PCR-based two-step DNA synthesis method for long gene sequences. *Nucleic Acids Res.*, **32**, e98.
- Sandhu,G.S., Aleff,R.A. and Kline,B.C. (1992) Dual asymmetric PCR: One-step construction of synthetic genes. *Biotechniques*, **12**, 14–16.
- Toung,L. and Dong,Q. (2004) Two-step total gene synthesis method. *Nucleic Acids Res.*, **32**, e59.
- Prodromou,C. and Pearl,L. (1992) Recursive PCR: A novel technique for total gene synthesis. *Protein Eng.*, **5**, 827–829.
- Stemmer,W.P., Cramer,A., Ha,K.D., Brennan,T.M. and Heyneker,H.L. (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene*, **164**, 49–53.
- Xiong,A.-S., Yao,Q.-H., Peng,R.-H., Duan,H., Li,X., Fan,H.-Q., Cheng,Z.-M. and Li,Y. (2006) PCR-based accurate synthesis of long DNA sequences. *Nat. Protoc.*, **1**, 791–797.
- Wu,G., Wolf,J.B., Ibrahim,A.F., Vadasz,S., Gunasinghe,M. and Freeland,S.J. (2006) Simplified gene synthesis: A one-step approach to PCR-based gene construction. *J. Biotech.*, **124**, 496–503.
- Kong,D.S., Carr,P.A., Chen,L., Zhang,S. and Jacobson,J.M. (2007) Parallel gene synthesis in a microfluidic device. *Nucleic Acids Res.*, **35**, e61.
- Huang,M.C., Ye,H., Kuan,Y.K., Li,M.-H. and Ying,J.Y. (2009) Integrated two-step gene synthesis in a microfluidic device. *Lab. Chip*, **9**, 276–285.
- Hoover,D.M. and Lubkowski,J. (2002) DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res.*, **30**, e43.
- Wu,G., Dress,L. and Freeland,S.J. (2007) Optimal encoding rules for synthetic genes: the need for a community effort. *Mol. Syst. Biol.*, **3**, 1–5.
- Lee,J.Y., Lim,H.-W., Yoo,S.-I., Zhang,B.-T. and Park,T.H. (2005) Efficient initial pool generation for weighted graph problems using parallel overlap assembly. *Lect. Notes Comp. Sci.*, **3384**, 215–223.
- Saleem,M., Kweon,M.-H., Johnson,J.J., Adhami,V.M., Elcheva,I., Khan,N., Hafeez,B.B., Bhat,K.M.R., Sarfaraz,S., Reagan-Shaw,S. *et al.* (2007) S100A4 accelerates tumorigenesis and invasion of human prostate cancer through the transcriptional regulation of matrix metalloproteinase 9. *Proc. Natl Acad. Sci. USA*, **103**, 14825–14830.
- SantaLucia,J. Jr and Hicks,D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, **33**, 415–440.
- Herrmann,M.G., Durtschi,J.D., Bromley,L.K., Wittwer,C.T. and Voelkerding,K.V. (2006) Amplicon DNA melting analysis for mutation scanning and genotyping: Cross-platform comparison of instruments and dyes. *Clin. Chem.*, **52**, 494–503.
- Wittwer,C.T., Reed,G.H., Gundry,C.N., Vanderstee,J.G. and Pryor,R.J. (2003) High-resolution genotyping by amplicon melting analysis using LCGreen. *Clin. Chem.*, **49**, 853–860.
- Giglio,S., Monis,P.T. and Saint,C.P. (2003) Demonstration of preferential binding of SYBR Green I to specific DNA fragments in real-time multiplex PCR. *Nucleic Acids Res.*, **31**, e136.
- Varga,A. and James,D. (2006) Real-time PCR and SYBR Green I melting curve analysis for the identification of *plum pox virus* strains C, EA, and W: effect of amplicon size, melt rate, and dye translocation. *J. Virol. Methods*, **132**, 146–153.
- Wittwer,C.T., Herrmann,M.G., Moss,A.A. and Rasmussen,R.P. (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. *Biotechniques*, **22**, 130–138.
- Ely,J.J., Reeves-Daniel,A., Campbell,M.L., Kohler,S. and Stone,W.H. (1998) Influence of magnesium ion concentration and PCR amplification conditions on cross-species PCR. *Biotechniques*, **25**, 38–40.
- von Ahsen,N., Wittwer,C.T. and Schütz,E. (2001) Oligonucleotide melting temperatures under PCR conditions: nearest-neighbor corrections for Mg^{2+} , deoxynucleotide triphosphate, and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. *Clin. Chem.*, **47**, 1956–1961.
- Rasmussen,J.P., Saint,C.P. and Monis,P.T. (2007) Use of DNA melting simulation software for *in silico* diagnostic assay design: Targeting regions with complex melting curves and confirmation by real-time PCR using intercalating dyes. *BMC Bioinformatics*, **8**, 107–118.
- Luo,R. and Zhang,D. (2007) Partial strands synthesizing leads to inevitable aborting and complicated products in consecutive polymerase chain reactions (PCRs). *Sci. China Ser. C-Life Sci.*, **50**, 548.
- Binkowski,B.F., Richmond,K.E., Kaysen,J., Sussman,M.R. and Belshaw,P.J. (2005) Correcting errors in synthetic DNA through consensus shuffling. *Nucleic Acids Res.*, **33**, e55.
- Fuhrmann,M., Oertel,W., Berthold,P. and Hegemann,P. (2005) Removal of mismatched bases from synthetic genes by enzymatic mismatch cleavage. *Nucleic Acids Res.*, **33**, e58.
- Bell,D.A. and DeMarini,D.M. (1991) Excessive cycling converts PCR products to random-length higher molecular weight fragments. *Nucleic Acids Res.*, **19**, 5079.
- Carr,P.A., Park,J.S., Lee,Y.J., Yu,T., Zhang,S. and Jacobson,J.M. (2004) Protein-mediated error correction for *de novo* DNA synthesis. *Nucleic Acids Res.*, **32**, e162.
- Don,R.H., Cox,P.T., Wainwright,B.J., Baker,K. and Mattick,J.S. (1991) ‘Touchdown’ PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res.*, **19**, 4008.

38. Ririe, K.M., Rasmussen, R.P. and Wittwer, C.T. (1997) Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Anal. Biochem.*, **245**, 154–160.
39. Blake, R.D., Bizzaro, J.W., Blake, J.D., Day, G.R., Delcourt, S.G., Knowles, J., Marx, K.A. and SantaLucia, J. Jr (1999) Statistical mechanical simulation of polymeric DNA melting with MELTSIM. *Bioinformatics*, **15**, 370–375.
40. Neuzil, P., Pipper, J. and Hsieh, T.M. (2006) Disposable real-time microPCR device: Lab-on-a-chip at a low cost. *Mol. BioSyst.*, **2**, 292–298.
41. Mamedov, T.G., Padhye, N.V., Viljoen, H. and Subramanian, A. (2007) Rational *de novo* gene synthesis by rapid polymerase chain assembly (PCA) and expression of endothelial protein-C and thrombin receptor genes. *J. Biotech.*, **131**, 379–387.
42. Arezi, B., Xing, W., Sorge, J.A. and Hogrefe, H.H. (2003) Amplification efficiency of thermostable DNA polymerase. *Anal. Biochem.*, **321**, 226–235.
43. Cherry, J., Nieuwenhuijsen, B.W., Kaftan, E.J., Kennedy, J.D. and Chanda, P.K. (2008) A modified method for PCR-directed gene synthesis from large number of overlapping oligodeoxyribonucleotides. *J. Biochem. Biophys. Methods*, **70**, 820–822.