

The Construction of a Short Gene by a Very Fast, Modified, and Simplified Gene Synthesis and the Analysis of Various Effects on This Synthesis

Mohammadreza Hajjari¹ and Behnaz Saffar^{1,2*}

¹Department of Genetics; Science Faculty of Shahrekord University; Iran. ²Biotechnology Research Department; Shahrekord University; Iran

ABSTRACT

We have constructed a short gene with a single step assembly PCR without any additional amplification primers with Taq and Pfu polymerases. Since the Taq polymerase is a common and conventional enzyme for PCR reactions, we have analyzed various effects on its efficiency. Eventually, we have been able to synthesize the gene in less than 40 minutes by Taq polymerase.

Key words: single step assembly PCR; efficiency; amplification primer; polymerase

INTRODUCTION

Since the first oligonucleotides assembly (Gupta et al., 1968; Kleppe et al., 1976; Crea et al., 1978; Goeddel et al., 1979) applied to gene synthesis, the technique has drawn much attention due to its extensive application in molecular biology and biotechnology. The examples of such applications to gene synthesis are **a)** the optimization of the DNA sequences for the heterologous expression in a suitable host (Wheeler et al., 1996; Milek et al., 2000; Manoj et al., 2004; Peng et al., 2006), **b)** the modification of the original sequences for specific purposes including elimination of sequences which instabilize RNA (Kang et al., 2004) **c)** establishing cryptic promoters (Bertorello et al., 1990) providing a short path to directed evolution (Xiong et al., 2007) and **d)** the investigation of transcriptional regulation (Krieg et al., 1991). One of the most critical situations which can be

addressed by chemical gene synthesis is the lacking of template DNA (Chang et al., 2002). Several methods including enzymatic ligation (Smith et al., 1982; Borodina et al., 2003), ForkI method (Hayden and Mandeki, 1988) and PCR-based strategies (Ciccarelli et al., 1991; Stemmer et al., 1995; Wooddell and Burgess, 1996; Gao et al., 2003; Young and Dong, 2004; Xiong et al., 2006) have already been applied to chemical gene synthesis.

In 1995, Stemmer et al described the assembly PCR technique, which was an attractive and relative low cost method in gene synthesis, as a technique in which small genes or gene fragments were synthesized. In this method short oligos which cover the entire desired gene duplex are selected. There is an overlapping between the successive oligos on sense and antisense strands. Conventionally, this method is performed in two steps. In the first step, the oligonucleotides is self-

*Author for correspondence: saffar_b@sku.ac.ir

assembled and primed by PCR. In the second step, in another round of PCR, two outside amplification primer are added and the gene is amplified (Stemmer et al., 1995).

Sequence errors are important limitations of assembly PCR. Much of these errors are introduced in second PCR reaction and sometimes screening of multiple colonies is required to select a colony without any errors (Baedeker and Schulz, 1999; Lin et al., 2002; Hu et al., 2006). This method was simplified by combining two PCR reactions into single reaction called "simplified gene synthesis". In this method, KOD polymerases perform better than *Pfu* and *Taq* polymerases which are conventionally used in assembly PCR. Despite their efforts to obtain full length 777 and 936 bp gene product by *Taq* or *Pfu* polymerases, no obvious full length was found. Then various parameters on efficiency of this technique were analyzed by KOD polymerases on the above-mentioned genes which resulted in full length 209 bp gene by *Taq* or *Pfu* but various parameters on its synthesis were not analyzed. In this study, a short gene was constructed with 195 bp length by modified SGS (Simplified Gene Synthesis). Then, its various effects such as concentration of oligonucleotides, primers and DNA polymerase,

the kind of DNA polymerase were analyzed. Finally, a very fast method for chemical gene synthesis with assembly PCR was established.

MATERIALS AND METHODS:

Gene and Oligo Design

The gene sequence which was selected for this study is *Synechococcus* PCC7942 *smtA* CDS (Gene accession number is X64585) that encodes metallothionein protein; its sequence was obtained from NCBI databases. In the modified method, Restriction sites of the enzymes "*BamH* I-*Nde*I" and *Hind* III were added to 5' end and 3' end of the gene's original sequence respectively so that it could be cloned in vectors such as pBluescript KS and pET 21. In fact, the overlapping oligo design was set upon the sequence that included the restriction site. This sequence was checked by Gene Runner program (version 3.05) in order to know whether the restriction sites for these enzymes were in the sequence or not. Then the composition of final sequence was checked by FREQSQ program (www.bioinfo.hku.hk/services/analyseq/cgi-bin/freqsq_in.pl). The characteristics of this sequence are shown in the Table1.

Table1 - Composition feature of the sequence.

53.33%	A+C
49.23%	A+G
45.64%	A+T
54.36%	G+C

The result showed that the sequence was a typical sequence. Then DNA work program was used to design the oligos in overlapping T_m which equaled 59° C. The oligos were changed in a way that their ends were only with G or C because *Taq* and *Pfu* polymerases work more efficiently with 3' end with G or C. six oligos were designed with a length ranging between 44 and 48 bp and

overlapping region between 15-19 bases. All oligonucleotides were checked by Gene Runner program for their GC content, hairpin loops, bulg loops, and dimers in 59°c. They were named Sm1 to Sm6. The oligos are listed in Table2. They were purchased from TAG Copenhagen A/S with RP-column purification and MALDI-TOF quality control.

Table 2 - Oligos sequence which were used in this study.

5'-AAA GGA TCC CAT ATG ACC TCA ACA ACG TTG GTC AAA TGC GCT TG-3'	sm1
5'-GGG ATC GAC GTT GCA GAG ACA GGG CTC ACA AGC GCA TTT GAC CAA C-3'	sm2
5'-CTC TGC AAC GTC GAT CCC AGC AAA GCG ATC GAT CGC AAC GGT CTG TAC-3'	sm3
5'-GGC CAT CGG CAC AGG CTT CGC TGC AGT ACA GAC CGT TGC GAT CG-3'	sm4
5'-CCT GTG CCG ATG GCC ACA CCG GTG GTA GCA AAG GCT GCG GCC ACA CC-3'	sm5
5'-AAA AAA AAG CTT TTA GCC GTG GCA GTT ACA GCC GGT GTG GCC GCA GCC-3'	sm6

Gene Synthesis

The solution of all the oligonucleotides was made with 100 μ M concentration by adding distilled ddH₂O. Then an assembly mix solution was made in which the concentration of each oligonucleotide was 1 μ M. By dilution of this mix, the desired concentration of each oligonucleotide in PCR reaction mix was obtained. In the first PCR reaction mix, a solution was made with 50 μ l volume including 5 μ l PCR buffer 10X, 1.5 μ l MgCl₂ (50mM) and 1 μ l dNTP (10mM). Then 1 μ l oligonucleotide and 2U *Taq* polymerase (Cinnagen) were added to the assembly mix. The final concentration of components in the PCR reaction mix was 1x (PCR buffer), 1.5mM(MgCl₂), 0.2mM(dNTP) and 20nM(oligonucleotides). For the second PCR reaction mix 0.5 μ l *Pfu* buffer10x with MgSO₄ (Fermantes), 1 μ l dNTP (10mM), 6U *Pfu* polymerase (Fermantes) and 1 μ l oligonucleotide were added. So the final concentration of each component in this reaction was 1x (*Pfu* buffer), 0.2mM (dNTP) and 20nM (oligonucleotides). In each reaction the first and the last oligos were used in excess amount (0.4 μ M). Each PCR reaction was carried out in 30 cycle of (i) denaturing in 94 $^{\circ}$ C for 30s, (ii) annealing in 57 $^{\circ}$ C for 30s and (iii) extension in 72 $^{\circ}$ C for 1 min. duration of final extension was 4 min.

Gel Electrophoresis and Sequencing of PCR Products:

The PCR products of reactions with *Taq* and *Pfu* polymerases were loaded on 1% agarose gel with 100 bp ladder and were sequenced by Macrogen Company.

RESULTS AND DISCUSSION

After observing the expected bands on the gel, they were sequenced. In the sequencing result there was no error in product of both reactions. Eventually various effects such as oligonucleotide concentration, amplification primers concentration, concentration of DNA polymerase, and the number of cycles were analyzed on construction of the gene.

Analysis of Oligonucleotide Concentration on Efficiency of Gene Synthesis

With the result of *Taq* and *Pfu* polymerases functioning well for short gene synthesis, various conditions for their function were analyzed so that these enzymes could be more applicable for short gene synthesis in single step assembly PCR. In accordance with the conventional assembly PCR, the equal molar of each oligonucleotide was used in each PCR reaction. Wu et al (2006) suggested that the optimal concentration of each oligonucleotide in single step assembly PCR was 10-25 nM.

However, this result was achieved by using KOD polymerases as well as the longer genes. In order to know whether using this range of concentration could be a rule for all of the polymerases and gene lengths, experiments were done with five different concentrations which four of them were out of this range. All of the reactions were performed in the above-mentioned conditions in material and method. The concentrations of the oligonucleotides in PCR reactions were 20, 50, 100 and 200 nM. Then PCR products were loaded on 1% agarose gel with 100bp ladder. Results (Figure1) showed that the optimal concentration of oligonucleotides in PCR reactions could be out of the range (10-25 nM). As the results show, in the concentration of 20, 50, and 100 nM, this method gave an acceptable result, but at 200 nM, the result was not desirable. To confirm the assumption that it was the extra concentration of oligos which affected the result, the oligos were used in PCR mix with 500 nM concentration. The result was not desirable too.

It appeared that the smears below the expected bands were the result of the oligos that were not bound to each other; not having eventually extended completely, they produced truncated genes with different lengths. Meanwhile, the smears upper than the expected positions were due to extra and nonspecific binding of gene fragments to each other and to oligos. It indicated that the amount of oligonucleotides that were added to the PCR mix might be too high.

This showed us that in single step assembly PCR with a definite PCR condition, the concentration of assembly oligos were critical, but depending on the conditions, these concentrations would be different.

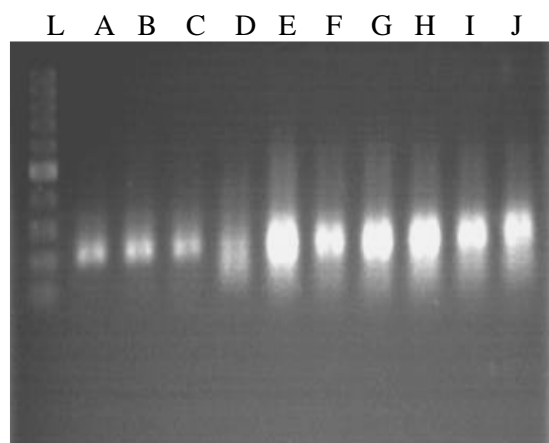


Figure 1 - Lanes A, B, C, D and E represent the PCR products by the oligonucleotides with 20, 50, 100, 200 and 500 nM respectively. The amount of *Taq* polymerases in these reactions was 2U. The lane F to J represents the PCR reaction products with 500nM and with 2, 2.5, 3, 4, 5 U *Taq* polymerase respectively (Ladder: 100bp, gel agarose 1%).

Analysis of *Taq* Polymerase Concentration on Efficiency of Gene Synthesis

The gene synthesis was done by 20 nM concentrations (an optimal concentration) of oligos with 2U, 4U and 6U *Taq* polymerase but no significant difference between PCR products bands on gel was observed (results not shown).

After observing the result for the 500 nM concentrations of oligos, to understand whether the concentration of *Taq* polymerase could affect the smears on gel electrophoresis or not, five different PCR reactions with different concentration of *Taq* polymerases was performed in another analysis. But results showed that there was no significant difference for this oligo concentration and smear bands persisted upon gel (Figure1). In the results of Wu et al (2006) experiment, these smear bands could be seen in PCR products with KOD polymerases, *Taq* and *Pfu* polymerase. The difference between these smear bands among the polymerases was that in reactions by *Taq* and *Pfu*, the smear bands were below the predicted size of the gene, but in the reactions by KOD polymerases they were both below and above it. These results could be attributed to the properties of KOD polymerases. The elongation rate and processivity of KOD polymerases were much more than the conventional polymerases (Takagi, et al., 1997). The oligos could not bind in reaction with *Taq* polymerase and remained intact because the

condition was not optimized for the PCR reaction. Meanwhile, in reaction with KOD polymerase, nonspecific longer genes were upon the gel in addition to truncated fragments because of the KOD polymerases properties and the non optimized PCR conditions. It is important to note that by performing the PCR with *Taq* and *Pfu* polymerase, and by increasing the concentration of oligos, the observed band on gel became nearer to the expected band. This occurrence was less poor by *Pfu* polymerase and it could be due to the smaller elongation rate as well as the processivity of *Pfu* compared to the *Taq* polymerase. Thus, it could be possible to construct longer genes by more concentration of *Taq* or *Pfu* polymerases or by modifying the PCR reaction condition.

Effect of Primer Concentration on Efficiency of Gene Synthesis

In this work, no additional external amplification primers were used. Instead, the first and the last oligonucleotides were employed in excess amount so that they could be used as amplification primers. The first and the last oligonucleotide could be designed in such a way that their ends functioned as sequences which include multiple restriction sites for different studies. In this part of the analysis, the oligos with 100nM concentration and 2U *Taq* polymerase were used.

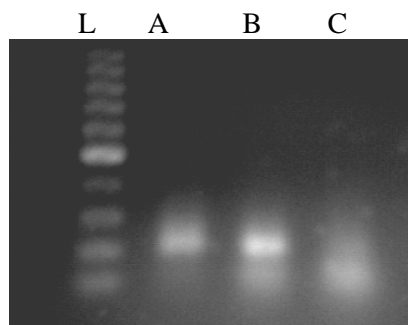


Figure 2 - Lanes A, B and C represent the one step assembly PCR products with 100nM oligonucleotide concentration and 2U *Taq* polymerase. Amplification primers:assembly oligos ratio is 25:1, 50:1 and 90:1 in lane A, B and C respectively (Ladder: 100bp, gel agarose 1%).

Eventually, the PCR reaction with the ratio of primer to oligos equal to 25:1, 50:1 and 90:1 was carried out.

Although the gene was constructed in the ratio of 50:1, it seemed that the excess amount of primers was not been served. Furthermore, in the ratio of 90:1, because of extra excess of the first and the last oligos, the yield of gene construction was low. In fact, in this concentration of primers the other oligos had no sufficient opportunity to bind together. By analyzing this result, it seemed that to achieve a more efficient one step assembly PCR product without any additional primers the concentration of the first and the last oligos should be optimized.

Very Fast Gene Synthesis by Single Step Assembly PCR

Finally, by various experiments (data not shown), a procedure was established in which gene synthesis could be carried out in a short time. In the PCR reaction mixes, the PCR buffer, $MgCl_2$ and dNTP concentration was the same as the above.

The amount of *Taq* polymerase was 6U and the number of the cycles decreased to 10. The results for the oligonucleotide with concentration of 20, 50, 100 and 200 nM are shown in Figure 3. Thus, a short gene was constructed in less than 40 minutes.

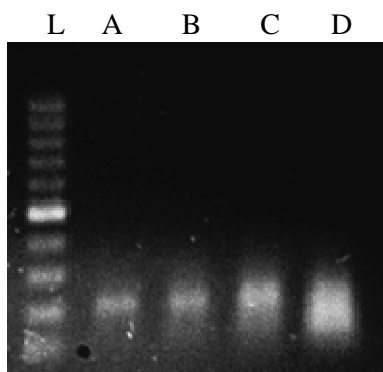


Figure 3 - Lanes A, B, C and D represent the very fast single step assembly PCR with 20, 50, 100 and 200 nM respectively. The concentration of the first and the last oligos as amplification primers was $0.4\mu M$ and the amount of *Taq* polymerase was 6 U (Ladder: 100bp, gel agarose 1%)

To analyze whether decreasing the number of cycles could affect the non-specific gene construction by oligonucleotides with 500nM concentration or not, this procedure was used for 500 nM oligo concentration; the results (Figure 4)

showed that in this condition There was not sufficient time to construct intact gene, and the excess oligos were shown in a smear manner above the gel.

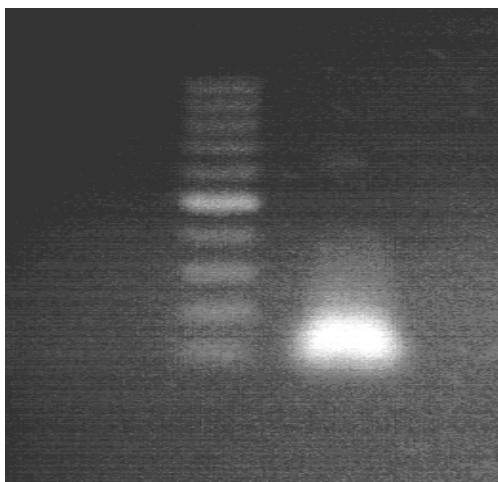


Figure 4 - The gene synthesis by 500nM oligo concentration, 0.4 μ M amplification primers and 6 U by very fast single step assembly PCR (Ladder:100bp, gel agarose 1%).

In short, in this study, a short gene was synthesized by single step assembly PCR with *Taq* and *Pfu* polymerases. This method was modified so that any additional external primers were not used. The first and the last oligonucleotides themselves were used as amplification primers. Moreover, since the *Taq* polymerase is a common and conventional enzyme for PCR reactions, some analysis were done to understand which parameters could affect the efficiency of the result. Finally, the short gene was constructed by an optimized procedure in less than 40 minutes.

By analyzing our results and comparing them with those of Wu et al (2006), with respect to the properties of KOD polymerases and their more elongation rate and processivity, it was confirmed that by modifying the PCR conditions such as oligonucleotides, polymerases and oligonucleotide concentration, by adjusting thermal cycle durations as well as the number of the cycles, the *Taq* and *Pfu* polymerases could be used for longer genes in single step assembly PCR.

Finally, It could be concluded that each of the above parameters, depending on the condition of the reaction (other parameters and the length of the

gene), could affect the yield of the simplified gene synthesis by *Taq* polymerase.

ACKNOWLEDGMENT

This study was supported by Genetic Department of Science Faculty at Shahrekord University in Iran. We thank Hossein Aliakbari for the Language correction of the manuscript.

REFERENCES

- Baedeker, M. and Schulz, G. E. (1999), Over expression of a designed 2.2 kb gene of eukaryotic phenylalanine ammonia-lyase in *Escherichia coli*. *FEBS Letter*, **457**, 57–60.
- Bertorello, A. M.; Hopfield, J. F.; Aperia, A. and Greengard, P. (1990), Inhibition by dopamine of (Na⁺+K⁺) ATP ase activity in neo striatal neurons through D1 and D2 dopamine receptor rsynergism. *Nature*, **347**, 386–388.
- Borodina, T. A.; Lehrach, H. and Soldatov, A.V. (2003), Ligation- Based synthesis of oligonucleotides with block structure. *Anal Biochem*, **318**, 309–313.

- Chang, B. S. W.; Kazmi, M. A. and Sakmar, T. P. (2002), Synthetic gene technology: applications to ancestral gene reconstruction and structure-function studies of receptors. *Method Enzymol*, **343**, 274–294.
- Ciccarelli, R. B.; Gunyuzlu, P.; Huang, J.; Scott, C. and Oakes, F.T. (1991), Construction of synthetic genes using PCR after automated DNA synthesis of their entire top and bottom strands. *Nucleic Acids Res*, **19**, 6007–6013.
- Crea, R.; Kraszewski, A.; Hirose, T. and Itakura, K. (1978), Chemical synthesis of genes for human insulin. *Proc Natl Acad Sci USA*, **75**, 5765–5769.
- 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.
- Goeddel, D. V.; Kleid, D. G.; Bolivar, F.; Heyneker, H. L.; Yansura, D. G.; Crea, R.; Hirose, T.; Kraszewski, A.; Itakura, K. and Riggs, A. D. (1979), Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *Proc Natl Acad Sci USA*, **76**, 106–110.
- Gupta, N. K.; Ohtsuka, E.; Sgaramella, V.; Buchi, H.; Kumar, A.; Weber, H. and Khorana, H.G. (1968), Studies on polynucleotides, 88. Enzymatic joining of chemically synthesized segments corresponding to the gene for alanine-tRNA. *Proc Natl Acad Sci USA*, **60**, 1338–1344.
- Hayden, M.A. and Mandecki, W. (1988), Gene synthesis by serial cloning of oligonucleotides. *DNA*, **7**, 571–577.
- Hu, S.; Li, L.; Qiao, J.; Guo, Y.; Cheng, L. and Liu, J. (2006), Codon optimization, expression, and characterization of an internalizing anti-ErbB2 single-chain antibody in *Pichia pastoris*. *Protein Express Purif*, **47**, 249–257.
- Kang, T. J.; Kang, K. H.; Kim, J.A.; Kwon, T. H.; Jang, Y. S. and Yang, M. S. (2004), High-level expression of the neutralizing epitope of porcine epidemic diarrhea virus by a tobacco mosaic virus-based vector. *Protein Express Purif*, **38**, 129–135.
- Kleppe, R.; Sekiya, T.; Loewen, P. C.; Kleppe, K.; Agarwal, K. L.; Buchi, H.; Besmer, P.; Caruthers, M. H.; Cashion, P. J.; Fridkin, M.; Jay, E.; Kumar, A.; Miller, R. C.; Minamoto, K.; Panet, A.; RajBhandary, U. L.; Ramamoorthy, B.; Sidorova, N.; Takeya, T.; van de Sande J. H. and Khorana, H. G. (1976), Total synthesis of the structural gene for the precursor of a tyrosine suppressor transfer RNA from *Escherichia coli*. 11. Enzymatic joining to form the total DNA duplex. *J Biol Chem*, **251**, 667–675.
- Krieg, R.; Stucka, R.; Clark, S., and Feldmann, H. (1991), The use of a synthetic tRNA gene as a novel approach to study in vivo transcription and chromatin structure in yeast. *Nucleic Acids Res*, **19**, 3849–3855.
- Lin, Y.; Cheng, G.; Wang, X. and Clark, T. G. (2002), The use of synthetic genes for the expression of ciliate proteins in heterologous systems. *Gene*, **288**, 85–94.
- Manoj, S.; Babiuk, L. A. and van Drunen Little-van den Hurk S. (2004), Approaches to enhance the efficacy of DNA vaccines. *Crc Cr Rev Cl Lab Sc*, **41**, 1–39.
- Milek, R. L.; Stunnenberg, H. G. and Konings, R. N. (2000), Assembly and expression Of a synthetic gene encoding the antigen Pfs48/45 of the human malaria Parasite Plasmodium falciparum in yeast. *Vaccine*, **18**, 1402–1411.
- Peng, R.H.; Yao, Q.H.; Xiong, A. S.; Cheng, Z. M. and Li, Y. (2006), Codon-modifications and an endoplasmic reticulum-targeting sequence additively enhance expression of an *Aspergillus phytase* gene in transgenic canola. *Plant Cell Rep*, **25**, 124–132.
- Smith, J.; Cook, E.; Fotheringham, I.; Pheby, S.; Derbyshire, R.; Eaton, M. A.; Doel, M.; Lilley, D. M.; Pardon, J. F.; Patel, T.; Lewis, H.; and Bell, L. D. (1982), Chemical synthesis and cloning of a gene for human beta-urogastrone. *Nucleic Acids Res*, **10**, 4467–4482.
- 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.
- Takagi, M.; Nishioka, M.; Kakihara, H.; Kitabayashi, M.; Inoue, H.; Kawakami, B.; Oka, M.; and Imanaka, T. (1997), Characterization of DNA polymerase from *Pyrococcus sp.* strain KOD1 and its application to PCR. *Appl Environ Microb*, **63**, 4504–4510.
- Wheeler, V. C.; Prodromou, C.; Pearl, L. H.; Williamson, R. and Coutelle, C. (1996), Synthesis of a modified gene encoding human ornithine transcarbamylase for Expression in mammalian mitochondrial and universal translation systems: A novel approach towards correction of a genetic defect. *Gene*, **169**, 251–255.
- Wooddell, C. I. and Burgess, R. R. (1996), Use of asymmetric PCR to generate long primers and single-stranded DNA for incorporating cross-linking analogs into specific sites in a DNA probe. *Genome Res*, **6**, 886–892.
- 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 Biotechnol*, **124**, 496–503.

- Xiong, A. S.; Yao, Q. H.; Peng, R. H.; Zhang, Z.; Xu, F.; Liu, J. G.; Han, P. L. and Chen, J. M. (2006), High level expression of a synthetic gene encoding *Peniophora lycii* phytase in methylotrophic yeast *Pichia pastoris*. *Appl Environ Microb*, **72**, 1039–1047.
- Xiong, A. S.; Peng, R. H.; Cheng, Z. M.; Li, Y.; Liu, J. G.; Zhuang, J.; Gao, F.; Xu, F.; Qiao, Y. S.; Zhang, Z.; Chen, J. M. and Yao, Q. H. (2007), Concurrent mutations in six amino acids in β -glucuronidase improves its thermostability. *Protein Eng Des Sel*, **20**, 319–325.
- Young, L. and Dong, Q. (2004), Two-step total gene synthesis method. *Nucleic Acids Res*, **32**, 59.

Received: May 11, 2009;
Revised: October 14, 2009;
Accepted: April 22, 2010.