American Journal of Current Microbiology

American Journal of Current Microbiology
http://ivyunion.org/index.php/ajcmicrob/index
Vo1. 1, Article ID 201300215, 8 pages

Research Article

pShuffle: A Plasmid for *in vitro* Evolution

D. Tischler*, P. Zwicker, R. Schwabe, J.A.D. Gröning, S.R. Kaschabek, M. Schlömann

Interdisciplinary Ecological Center, TU Bergakademie Freiberg, Freiberg, Germany

Abstract

Multi-gene shuffling is a powerful method used to combine and optimize attributes of various proteins. Here we report on the design and construction of the plasmid "pShuffle" which is suited for a variety of *in vitro* DNA-recombination techniques. The multiple cloning site (MCS) of pShuffle was designed to allow for the cloning of genes as well as their expression under control of either a lac- or a T7-promoter. As a specific feature, this MCS allows for the fusion of special linker sequences to both ends of cloned genes. After subsequent DNA-recombination steps, these linkers facilitate reamplification of generated gene variants, and thus may be used to construct clone libraries for activity screenings. The suitability of pShuffle for multi-gene shuffling applications was further shown with a set of styrene monooxygenase genes originating from proteo- and actinobacteria.

Keywords: DNA-shuffling; Enzyme evolution; Flavin monooxygenase; Multi-gene shuffling; DNA-recombination

Peer Reviewer: Guliy Olga Ivanovna, PhD, Institute of Biochemistry & Physiology of Plants & Microorganisms, Russian Academy of Sciences, Russian Federation; K M Singh, PhD, Dept of Animal Biotechnology, Anand Agricultural University (AAU), India

Executive Academic Editor: Sihua Peng, Shanghai Ocean University

Received: August 28, 2013; **Accepted**: October 19, 2013; **Published: November** 4, 2013

Competing Interests: The authors have declared that no competing interests exist.

*Correspondence to: Dirk Tischler, Interdisciplinary Ecological Center, Environmental Microbiology Group, TU Bergakademie Freiberg, Leipziger Str. 29, 09599 Freiberg, Germany; Email: dirk-tischler@email.de

Introduction

Flavin-dependent monooxygenases catalyze highly regio- and enantioselective oxofunctionalization reactions [1]. However, so far virtually none of these enzymes have proven suitable for industrially relevant biotechnological processes, a fact which may be attributed to a number of reasons, including: *i)* narrow substrate range (typical for hydroxylases), *ii)* cofactor dependence, *iii)* the phenomenon of uncoupling (as in Baeyer-Villiger monooxygenases and styrene

monooxygenase), *iv)* multi-component nature (as in many oxygenases), and *v)* insufficient specific activity and/or stability.

Among flavin monooxygenases the class of styrene monooxygenases (SMOs) has shown the most promise as biocatalysts for the synthesis of fine chemicals, but limited available structural information currently restricts the directed evolution of such enzymes by rational design [2]. However, an approach of rational design in combination with computational modeling for substrate docking into the active site in order to predict the involved residues has shown some success [3]. Error-prone polymerase chain reaction (PCR)-based techniques have also been used as an undirected *in vitro* method to improve specific activity of an SMO originating from *Pseudomonas* [4]. In view to these findings, methods of undirected evolution or DNA-recombination seem to hold the most promise for biocatalyst optimization without requiring additional structural studies, since these techniques produce higher sequence diversity among mutants than previously used methods [5-8].

Two types of bacterial SMOs can be distinguished with respect to their polypeptide chains, specific activity, and oxygenation efficiency [2, 9]. The SMO encoded by *styA*/*styB* is typical for *Pseudomonas* species [10], while that encoded by *styA1*/*styA2B* was first found in the actinobacterium *Rhodococcus opacus*1CP [11]. Catalytic features of both bacterial SMOs could be potentially valuable for improving biocatalyst performance when combined in a novel variant.

Multi-gene shuffling is a useful tool that combines features of homologous proteins, and is used currently to accomplish protein evolution. However, as microgram quantities of the corresponding gene templates are required for this technique, it is more accessible by means of plasmid propagation than by PCR [8]. Subjecting the complete gene length to the process of homologous recombination and achieving a maximum number of cloned and expressed variants are also required. Therefore, simplifying the reamplification of mutated genes and their subsequent cloning would be advantageous. The presented study describes the construction and evaluation of a specific cloning plasmid “pShuffle”, a genetic tool that meets the requirements for improving these techniques.

Materials and Methods

Cultivation and cloning. The host *E. coli* DH5α was used for cloning, plasmid propagation, and pShuffle-based gene expression. It was cultivated aerobically in LB medium at 37°C, unless stated otherwise [12]. Medium was supplemented with ampicillin (100 µg ml^{-1}) for plasmid propagation; isopropyl β-D-1-thiogalacto-pyranoside (IPTG; 0 or 2 mM) was used for gene expression.

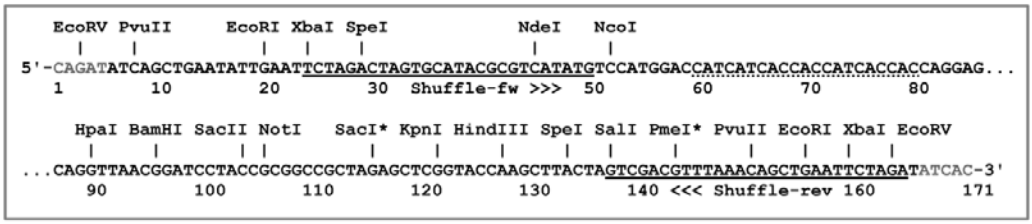
The restriction-digestion of plasmid DNA, agarose gel electrophoresis, plasmid DNA extraction and purification, cloning procedures, preparation of competent cells and heat-shock transformation were all performed as described previously [12-14].

Polymerase chain reaction (PCR) was performed using DreamTaq DNA polymerase and an appropriate quantity of template DNA according to manufacturer’s instructions. Gene-specific primers were used to amplify *styA* from *Pseudomonas fluorescens* DSM6290 [10] (SA-fw: 5’-CATATGAAAAAGCGTATCG-3’ and SA-rev 5’-GCGGCCGCTCAGGCCGCGATAG-3’).

The pShuffle-inserts or products obtained from DNA-shuffling were amplified by using linker-specific primers (Shuffle-fw: 5’-CCACTGTCTAGACTAGTGCATACGCGT CATATG-3’ and Shuffle-rev:

5'-GCAGCTCTAGAATTCAGCTGTTTAAAC GTCGAC-3'). PCR products obtained were purified and ligated in either pJET1.2 (Thermo Scientific) or pShuffle_Lac.

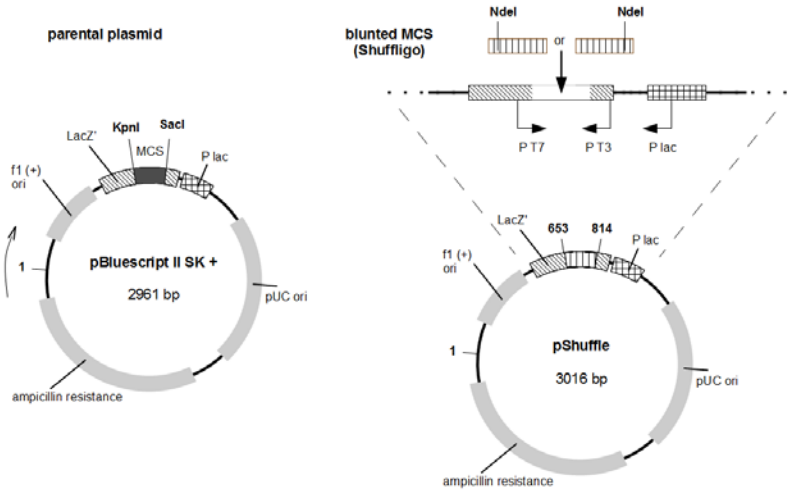
The plasmids pSRoA1_P01 and pSRoA2B_P01 served as sources for *styA1* and *styA2B*, respectively [11].Genes were excised via the restriction sites NdeI and NotI and subsequently ligated into similarly treated pShuffle_Lac, yielding the constructs pShuffle-styA1 and pShuffle-styA2B, respectively.



Scheme 1 Shuffligo (171-bp) harboring the polylinker for pShuffle

Bases are numbered and restriction sites are indicated at the first base of their recognition motif. EcoRV-sites provide the cleavage sites to generate a blunt-end oligonucleotide (black) representing the MCS of pShuffle-variants. The SacI* and PmeI* sites may also be cleaved using Eco53kI and DraI, respectively. The primer-binding sites for the reamplification after shuffling are underlined. The start-codon of the NdeI-site can be used in combination with a His_6-tag-coding sequence (dotted underline) for the production of His_6-tag fusion proteins.

pShuffle construction. Sequence analysis was performed with the online-tool NEBcutter (version 2.0; New England Biolabs). A list of suitable restriction sites was generated and used to design Shuffligo (Scheme 1), which contained the MCS for the constructed plasmids. The oligonucleotide was synthesized and delivered as plasmid pCR2.1-Shuffligo (Eurofins MWG Operon).



Scheme 2 Construction of two pShuffle variants

Plasmids are numbered clockwise and ORFs are counter-clockwise in orientation. Features displayed: origin of single-stranded DNA replication (f1 (+) ori), sequence coding for β-galactosidase α-fragment (LacZ'), the

T7 and T3 promoter transcription initiation sites (P T7 and P T3), the lac promoter (P lac), the origin of replication from pBR322 (pUC ori), and the ORF coding for ampicillin resistance. The MCS was excised from the parental plasmid **(A)** and the MCS (Shuffligo; dashed) derived from pCR2.1-Shuffligo by EcoRV restriction digestion was inserted to yield pShuffle **(B)**. Depending on the orientation, ligation may lead to pShuffle_Lac (NdeI neighbored to P lac) or pShuffle_T7 (NdeI neighbored to P T7).

The phagemid pBluescript II SK+ served as the parental plasmid for the construction of pShuffle (Scheme 2). Its polylinker region was excised by KpnI/SacI-double digest and the obtained ~2,800-bp fragment was blunted (blunting enzyme, pJET1.2 cloning kit, Thermo Scientific) and served as the core for the novel MCS derived from Shuffligo. EcoRV was used to excise the MCS out of pCR2.1-Shuffligo, and attempts to ligate this MCS into linearized pBluescript yields two variants of pShuffle (orientation is given in Scheme 2). The products obtained from ligation of the MCS and the pBluescript fragment were transformed into *E. coli* DH5α and propagated. Plasmids of several clones were controlled by restriction digestion and sequencing.

DNA-Shuffling. In order to prepare sufficient amounts of template DNA, pShuffle-derivatives (pShuffle-styA1, pShuffle-styA2B) were propagated. Plasmid extraction (NucleoBond AX100, Macherey-Nagel, Germany) was performed on a total of 25ml overnight culture according to the manufacturer's instructions. Template genes were prepared by PvuII-digest of these plasmid preparations. DNA concentration of individual steps was spectrophotometrically controlled (NanoDrop; PeqLab) with nuclease-free water as the standard. Fragmentation of templates was achieved by a controlled DNaseI treatment in the corresponding buffer. Equal amounts of template DNA of both genes with shuffle-sites (1 to 10 µg) were mixed and subjected to the fragmentation procedure. DNaseI activity (25 mU per 1 µg DNA) was adjusted according to the amount of total template DNA; various incubation temperatures (4 to 25°C) and reaction times were also tested. The reaction was initiated by the addition of enzyme and terminated by the addition of EDTA (final concentration 50 µM) and heat-inactivation (96°C, 12 min). The total fragments obtained and the progress of the reaction were controlled by agarose gel electrophoresis (2% gel). Gene fragments were purified with the Easy Pure DNA Purification Kit from Biozym (~30% yield of applied DNA). About 0.3 to 7.2 ng μl^{-1} of the template DNA was subjected to a reassembling PCR omitting additional primers. Various reaction conditions for reassembly were tested and controlled by agarose gel electrophoresis. In general, all reactions were performed with DreamTaq DNA polymerase (2.5 U) in the corresponding buffer (Thermo Scientific) including DMSO (up to 7.5 % v/v) in a total volume of 100 µl. The temperature program included a denaturation step at 95°C for 5 min, then 34 cycles of the following three steps: *i)* denaturation at 95°C for 30 s, *ii)* annealing at various temperatures for 30 s; and *iii)* elongation at 72°C for 30 s (additional 5 s per cycle), followed by a final elongation step at 72°C for 5 min. The following annealing temperatures were applied: constant 18°C; 19°C + 0.3°C per cycle (up to 29°C); constant 48°C; or 49°C + 0.3°C per cycle (up to 59°C). After reassembling, a second stage of PCR was performed as described above with the linker-specific primers. Products obtained were controlled for size, gel-purified, and subsequently used to generate a clone library (pShuffle-A1-SR1). Amplicons of a suitable size (StyA1_R1) as well as the plasmid pShuffle were subjected to a PmeI/NdeI-double

digest, purified, mixed and subjected to a ligation reaction with T4 DNA ligase. After transformation into competent *E. coli* DH5α, clones were plated onto LB medium containing antibiotics and colonies were screened for SMO activity. Several plasmids from such clones were isolated and sequenced.

Results and Discussion

SMO genes are frequently detected in the genomes of GC-rich microorganisms from different phyla [9] and as a consequence show high sequence diversity. This makes it difficult to apply DNA-shuffling methods to these genes, since a certain degree of sequence similarity is required in most cases [6, 7]. Additionally, common cloning and/or expression plasmids were found to be unsuitable to fulfill all requirements for DNA-shuffling. Cloning can be done with almost every commercially available plasmid, however, but to excise cloned genes by a simple restriction endonuclease treatment and to simultaneously provide genes with defined linker sequences (shuffle-sites) seems unlikely to succeed for a variety of genes. As described in this work, using such linkers adjacent to both sites of genes allows for DNA-recombination of whole-length genes and their subsequent reamplification, while maintaining defined restriction sites for further cloning strategies.

In order to generate a suitable plasmid that facilitates cloning of numerous genes from different microorganisms (e.g. SMOs) and that allows the preparation of genes for various DNA recombination experiments, the following steps were successfully performed: *in silico* sequence analysis, conventional cloning approaches yielding pShuffle variants, and a DNA-shuffling procedure [6] for evaluation of the technique.

The sequence analysis revealed numerous restriction enzymes which were found to be rare cutters, allowing us to design a novel polylinker sequence suitable for cloning various genes from different microorganisms (Scheme 1). In addition to these restriction sites, linker sequences which can serve as PCR-primer binding sites (Shuffle_fw, Shuffle_rev) and restriction motifs that allow excising inserts by means of a single restriction endonuclease (EcoRI, PvuII, XbaI) were included. The core of the novel plasmid included the phagemid pBluescript II SK+ (Scheme 2). The novel polylinker could be introduced in two directions into the pBluescript-fragment. Construction of both variants was performed and controlled as described above. Depending on the orientation of the NdeI-site of the inserted MCS, which can be located proximate to the lac- or to the T7-promoter, constructs were designated pShuffle_Lac and pShuffle_T7, respectively.

In order to demonstrate the functionality of pShuffle, the SMO-genes *styA* (*Pseudomonas fluorescens* DSM6290), *styA1* (*Rhodococcus opacus* 1CP) and *styA2B* (*R. opacus* 1CP) were cloned into pShuffle_Lac. Therefore, *styA* was amplified by PCR from strain DSM6290. A product of correct size (~1,248bp) was obtained at an annealing temperature of 50°C in the presence of bovine serum albumin (0.2 mg ml^{-1}) and DMSO (7.5% v/v). The product was subsequently gel-purified, blunted, ligated into pJET1.2, and propagated. The *styA*-gene was obtained from the pJET1.2-derivative by NdeI/NotI double restriction digestion and cloned into similarly treated pShuffle_Lac, yielding the construct pShuffle-styA. Correct insertion was verified by restriction digest and sequencing (not shown). The plasmids pShuffle-styA1 and pShuffle-styA2B were successfully constructed as described above. They were propagated, prepared, and the correct size of the inserts was controlled by PCR. Defined amounts of these plasmid preparations were later used as parental gene sources for shuffling approaches. However, the successful

cloning showed the suitability of pShuffle as a cloning plasmid. Besides the previously mentioned restriction endonucleases, additional enzymes with recognition motifs in the MCS of pShuffle were tested and were found to function as predicted.

The applicability of pShuffle_Lac as an expression system was shown by means of transformants of pShuffle-styA. Clones were plated on LB medium and incubated at 30°C over night. Interestingly, the color of colonies turned dark blue, indicating the presence of an active and soluble SMO. The presence of an active tryptophanase in *E. coli*, low concentrations of indole present in the medium and an active SMO produced from the same host allowed for the conversion of indole to indole oxide, which spontaneously dimerizes to yield the water-insoluble and dark blue dye indigo [11]. The inducer IPTG and the substrate indole (1 mM) were plated onto agar plates to enhance the SMO gene expression and to improve colorimetric screening, respectively. Again the formation of indigo was observed during cultivation of *E. coli* DH5α (pShuffle-styA), indicating SMO activity and that pShuffle was suited for gene expression. Additionally, it shows that pShuffle-SMO constructs can be subjected to activity screening approaches, which help to simplify the procedure of directed evolution. Even in the absence of the inducer IPTG, the blue insoluble dye was formed, which indicates a certain extent of basal expression of the monooxygenase. Like the parental plasmid pBluescript, the construct pShuffle_Lac does not possess a tightly regulated lac-promoter, yielding some basal expression of inserted genes.

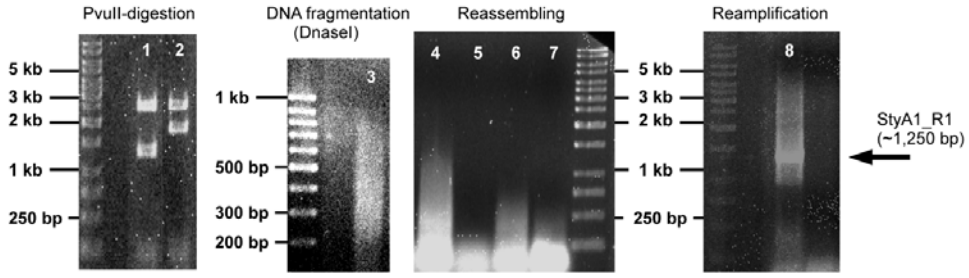


Figure 1 Multi-gene shuffling of two SMOs according Stemmer [6]

The SMO-genes *styA1* (~1,221bp) and *styA2B* (~1,722bp) were cloned into pShuffle_Lac and propagated. Genes plus linker sequences were then excised by PvuII-digestion (Samples1 and 2), gel-purified, and subjected to controlled DNaseI treatment (3). 250 to 350 bp-fragments obtained were purified via gel electrophoresis and served as template in the reassembling reaction. The latter represents a primer-less PCR in which fragments bind to each other and genes can be reconstituted. Four different conditions were assayed (4 to 8), from which one (4) yielded products up to a desired size of about 1,250 bp, which served as a template for the reamplification process. Products obtained from this process (8) were further gel-purified, and represent the SMO-library StyA1_R1.

A protocol for DNA-shuffling incorporating pShuffle as a supporting tool was established [6]. The above-mentioned pShuffle derivatives harboring *styA1* or *styA2B* served as the starting point of a shuffling approach. An overview of the procedure and products obtained from each step is given in Figure 1.

Sufficient amounts of template DNA were prepared *via* plasmid propagation from the pShuffle derivatives. The overall yield was

about 280 µg purified plasmid DNA from each preparation. Template genes were prepared using a single restriction enzyme (PvuII). The ability to generate large quantities of template DNA for shuffling approaches is one of the major advantages of the pShuffle system. In addition, genes of interest were obtained with adjacent shuffle sites (linker sequences) (Scheme 1), which was of benefit later during the procedure (e.g. reamplification, or cloning). The DNaseI-fragmentation of templates was performed as described above. A fragment size range of 50 to 300bp is required to provide proper fragments for the next step of DNA-shuffling. Suitable fragmentation was achieved during the DNaseI treatment at 4°C for 2 min (mean fragment size of 250bp; Figure 1).These gene fragments were purified and subsequently subjected to reassembling PCR. Different reaction conditions were tested and controlled. In almost all cases the majority of products obtained had a size of the applied fragments or slightly higher. In a few cases, only a weak smear around the expected size of about 1,221bp was observed. In this case (Figure 1, Sample 4) higher annealing temperatures (constant 48°C or 49°C plus ramp) and higher amounts of template DNA yielded better results. The reassembling product obtained was used as the template for the reamplification PCR including the linker-specific primer pair Shuffle_fw/Shuffle_rev (annealing temperature 56°C). The products obtained were analyzed (Figure 1) and subsequently used to generate a clone library (pShuffle-A1-SR1). Here the linkers allowed for the simple reamplification of mutated genes, even if recombination events occurred at the initial or last bases of genes. They also allowed the start codon to be maintained and subsequently cloned into pShuffle_Lac (*via* PmeI and NdeI) for screening purposes. Several plasmids derived from such clones were controlled by sequencing. It was confirmed that SMO genes were reconstituted from scratch, and that DNA-shuffling using the tool pShuffle was successfully performed.

Concluding Remarks

Two pShuffle variants were constructed and evaluated. The MCS of pShuffle allows the cloning of numerous genes and the simple fusion of linker sequences to both sites of genes prior to a shuffling application. This allows for the production of large quantities of template DNA by plasmid propagation. A gene plus adjacent linkers can be excised by a single restriction endonuclease, since three restriction sites are present at each site (EcoRI, PvuII, and XbaI; Scheme 1). Linker sequences can be used following shuffling for a simple reamplification of randomly generated gene variants and for subsequent cloning into plasmids to construct clone libraries for activity screening. Furthermore, pShuffle enables the shuffling of complete genes from start-to-stop codons. This is a significant improvement over previous techniques, as it allows the covering of as much sequence space as possible in subsequent DNA-recombination processes [5, 7]; this should also prove advantageous for techniques such as error-prone PCR. Finally, these constructs may serve as expression plasmids, under control of either the lac- or the T7-promoter. The addition of a His_6-encoding region may also significantly facilitate purification of shuffled protein-variants. Thus, pShuffle can be considered a promising tool for *in vitro* evolution of proteins.

Protocols for *in vitro* DNA-recombination provide emerging tools to mimic natural evolution within a short timescale, and may be used in combination with highly specific screening approaches in a more directed fashion [5-8]. The plasmid pShuffle serves as an adaptive tool for DNA-shuffling procedures,

especially for the directed evolution of monooxygenases. In addition to the SMOs focused on in this study, *in silico* sequence analysis revealed the applicability of pShuffle and its MCS for cloning and for recombination of additional flavin monooxygenases like hydroxylases and Baeyer-Villiger monooxygenases with interesting and biocatalytically relevant properties [1].

Acknowledgement

The project was supported by the Saxon Ministry for Environment, Agriculture, and Geology (LfULG-1771508003). Dirk Tischler was supported by a predoctoral fellowship from the Deutsche Bundesstiftung Umwelt (DBU). The authors declare no conflict of interest. We thank Eliot Morrison (FU Berlin) for carefully reading and discussing the manuscript.

References

1. van Berkel WJH, Kamerbeek NM, Fraaije MW. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. ***J Biotechnol***. 2006, 124:670-689
2. Montersino S, Tischler D, Gassner GT van Berkel, WJH. Catalytic and structural features of flavoprotein hydroxylases and epoxidases. ***Adv Synth Catal***. 2011, 353:2301-2319
3. Qaed AA, Lin H, Tang D-F, Wu Z-L. Rational design of styrene monooxygenase mutants with altered substrate preference. ***Biotechnol Lett***. 2010, 33:611-616
4. Gursky L, Nikodinovic-Runic J, Feenstra K, O'Connor K. In vitro evolution of styrene monooxygenase from *Pseudomonas putida* CA-3 for improved epoxide synthesis. ***Appl Microbiol Biotechnol***. 2010, 85:995-1004
5. Crameri A, Raillard S-A, Bermudez E, Stemmer WPC. DNA shuffling of a family of genes from diverse species accelerates directed evolution. ***Nature***. 1998, 391:288-291
6. Stemmer WPC. DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. ***Proc Natl Acad Sci USA***.1994, 91:10747-10751
7. Yuan L, Kurek I, English J, Keenan R. Laboratory-directed protein evolution. ***Microbiol Mol Biol Rev***. 2005, 69:373-392
8. Zhao H, Arnold FH. Optimization of DNA shuffling for high fidelity recombination. ***Nucleic Acids Res***. 1997, 25:1307-1308
9. Tischler D, Gröning JAD, Kaschabek SR, Schlömann M. One-component styrene monooxygenases: an evolutionary view on a rare class of flavoproteins. ***Appl Biochem Biotechnol***. 2012, 167:931-944
10. Beltrametti F, Marconi AM, Bestetti G, Galli E, Ruzzi M, Zennaro E. Sequencing and functional analysis of styrene catabolism genes from *Pseudomonas fluorescens* ST. ***Appl Environ Microbiol***.1997, 63:2232-2239
11. Tischler D, Eulberg D, Lakner S, Kaschabek SR, van Berkel WJH, Schlömann M. Identification of a novel self-sufficient styrene monooxygenase from *Rhodococcus opacus* 1CP. ***J Bacteriol***. 2009, 191:4996-5009
12. Sambrook J, Fritsch E, Maniatis T. Molecular cloning: a Laboratory Manual, 3rd edn. ***Cold Spring Harbor Laboratory Press***, Cold Spring Harbor, N.Y., 2001
13. Inoue H, Nojima H, Okayama H. High efficiency transformation of *Escherichia coli* with plasmids. ***Gene.*** 1990, 96:23-28
14. Marchuk D, Drumm M, Saulino A, Collins FS. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. ***Nucleic Acids Res***. 1991, 19:1154