Biotechnol. Bioprocess Eng. **2002**, 7: 311-315

In Vitro Combinatorial Mutagenesis of the 65th and 222nd Positions of the Green Fluorescent Protein of *Aequarea victoria*

Hideo Nakano*, Reiko Okumura, Chinatsu Goto, and Tsuneo Yamane

Laboratory of Molecular Biotechnology, Graduate School of Biological and Agricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan

Abstract By the *in vitro* combinatorial mutagenesis, which is a sequential reaction of PCR mutagenesis and *in vitro* coupled transcription/translation with *Escherichia coli* S30 extract, S65 and E222 of green fluorescent protein of *Aequarea victoria* were comprehensively changed to all possible combinations of amino acids, thus totally 400 mutant (including a wild type) proteins were simultaneously produced and their fluorescent properties were analyzed. Although a few mutations had been reported so far at the 222nd position, replacement E222 to all other19 amino acids gave fluorescent signal to the mutants by changing Ser 65 to Ala together. Among the mutants, replacement to G, A, S, Q, H and C gave relatively high fluorescence. The *in vitro* combinatorial mutagenesis, therefore, has been proved valuable for comprehensive structure-function studies of proteins.

Keywords: cell-free protein synthesis, polymerase chain reaction, green fluorescent protein, combinatorial mutation

INTRODUCTION

Site-specific introduction of mutations into a protein of which three-dimensional structure is known is a powerful tool especially to investigate the function of each residue in proteins scientifically. In order to engineer proteins to fit a particular purpose, such as an increase in thermal stability, activity for a specific substrate, etc, however, the site-directed method has not been so effective, although a number of successes have been reported so far. This is because the current understanding of structure, function, and folding of proteins is not enough for precise prediction of the effect of mutations, and making a mutant protein with even one amino acid replacement is actually a labor-intensive work. It needs an *in vitro* mutation of a target gene on an expression vector by usually single-stranded template or by PCR, transformation of *E. coli* cells with the mutated ones, cultivation of several candidates, selection by DNA sequencing, and finally expression in appropriate host cells. Since the whole step is necessary for each mutant protein, relatively a long time is needed to prepare a mutant protein. Therefore to get a successful mutant is a laborious work.

The most excellent method to reduce the task for the construction of site-specific mutant proteins is, as far as we know, the direct combination of *in vitro* mutagenesis mediated by PCR, overlap extension to reconstruct the whole DNA sequence, and *in vitro* coupled transcription/translation. By using thus reconstituted template as bulk, the labor-intensive cloning and sequencing steps can be omitted. The scanning introduction up to 19 amino acid replacements at a unique site has been reported by Burks *et al.* [1]. They analyzed the binding pocket of an antibody by means of this method intensively and thoroughly.

We have modified their method to enable the introduction of multiple mutations into a protein in a combinatorial manner, and successfully changed a substrate specificity of *Burkholderia cepacia* KWI-56 lipase by the combinatorial introduction of mutation into the substrate binding site [2].

Here to analyze the relationship between S65 in the fluorophore and E222 of green fluorescent protein (GFP) of *Aquarea victoria,* which is fluorescent without any cofactor and has become an inevitable tool for biological and biotechnological research [3], totally 400 mutant GFPs that includes one wild type were produced *in vitro,* and their fluorescent characters were analyzed. Finally, the potential application of this method especially for protein design is discussed.

MATERIALS AND METHODS

The outline of the *in vitro* combinatorial mutagenesis is shown in Fig. 1. Plasmid pRSET-GFP-1, which was constructed by inserting the green fluorescent protein gene of pGFP (Clonetech, USA) into pRSETb vector (Invitrogen, USA), was used as the source of the *gfp*

* **Corresponding author**
Tel: +81-52-789-4143 Fax: +81-52-789-4145
e-mail: hnakano@agr.nagoya-u.ac.jp

gene [4]. To decrease the contamination of the wild type gene in mutant, template DNA was first digested by restriction enzymes into two fragments. Each fragment was purified from 1% agarose gel by the GeneElute agarose spin column (Sigma, USA). In Step 1, three DNA fragments, f1, f2, and f3 were independently amplified by primer sets, F1T1-S65overlap, S65X-E22overlap, and E222X-R1T1, respectively. 'S65X' means primers replacing 65th Ser to other 19 amino acids, and 'E222X' means E222 to other 19 amino acids. The PCR was done in 20 μL of 1×ExTaq DNA polymerase buffer containing 0.2 μM each primer, 0.05 U/μL ExTaq DNA polymerase (Takara Shuzo, Japan), 0.2 mM dNTPs, in the following temperature sequence; 94°C for 2 min, 25 cycles of 95°C for 30 sec, 55°C for 30 sec, and 72°C for 2 min, followed by additional extension at 72°C for 10 min, using GeneAmp PCR system 2400(PerkinElmer). After estimation of each amplification by agarose gel electrophoresis, in Step2, totally 2 μL of the three fragments covering the full *gfp* gene, T7 promoter, and its terminator was joined by overlap extension PCR as follows. First, the fragments was joined in 25 μL of 1× ExTaq DNA polymerase buffer containing 0.05 U/μL ExTaq DNA polymerase and 0.2 mM dNTPs by the following temperature sequence: 94°C for 2 min, 5 cycles of 95°C for 30 sec, 46°C for 30 sec, and 72°C for 2 min. Then after addition of 2 μM T1 primer, in Step3 exactly the same temperature sequence of step 1 was carried out again. The reconstitution of the fragment was confirmed by agarose gel electrophoreses. In addition, DNA sequence of the fragment thus amplified was analyzed directly from the amplified fragment as template, confirming the expected mutation.

Then the amplified fragment was served as template for coupled transcription/translation using *E. coli* S30 extract as described previously with minor modifications [5]. Typically the reaction mixture of the coupled transcription/translation contained 56.4 mM Tris-acetate (pH 7.4), 1.76 mM dithiothreitol, 1.22 mM ATP, 0.85 mM GTP, CTP, and UTP, 0.7 mM 20 amino acids, 40 mM creatine phosphate, 4% (w/v) polyethylene glycol 6000, 34.6 μg/mL folinic acid, 0.17 mg/mL *E. coli* tRNA, 35.9 mM ammonium acetate, 0.15 mg/mL creatine kinase, 6.7 mM $Mg(OAc)_2$, 120 mM KOAc, 10 μg/mL T7 RNA polymerase, 10% of the PCR reaction mixture, and 28.3% S30 extract.

For the direct measurement of *in vitro* synthesized GFP mutants, the reaction mixture of *in vitro* transcription/translation incubated for 2 h at 37°C was diluted 20-fold with 100 mM sodium phosphate buffer (pH 7.0), and its fluorescence was measured by using a spectroflurometer F4500 (Hitachi, Japan).

To purify thus *in vitro* synthesized GFP, 200 μL of reaction was incubated for 1 h at 4°C after the addition of 20 μL of Ni chelate beads solution (Ni-NTA Magnetic Agarose Beads, Qiagen, USA). Then it was washed twice with 200 mL of 50 mM Na-phosphate buffer (pH 6.3), once with 200 μL of the above buffer additionally containing 10 mM imidazole and eluted with 50 μL of the buffer containing 250 mM imidazole.

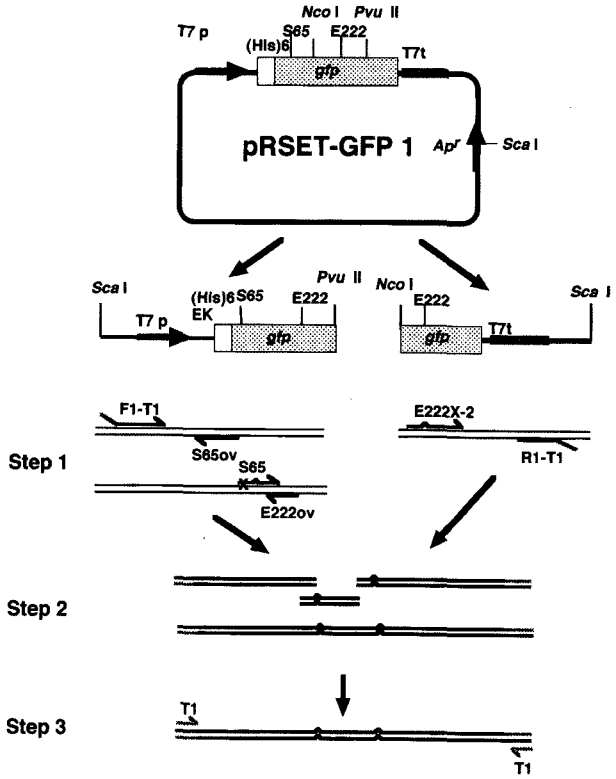

Fig. 1. Schematic representation of construction of mutant genes by PCR mutagenesis and overlap extension T7p, T7t, (His)6, and Ap^r represent T7promotor, T7 terminator, hexahistidine tag, and ß-lactamase, respectively.

RESULTS AND DISCUSSION

Methodological Improvements in the *In Vitro* Combinatorial Mutagenesis

Elimination of the amplification of the original sequence in overlap extension PCR is important to assure reliability of changes in phenotype obtained by the *in vitro* method, because no cloning step is included. In previous reports [1,2], gel purification of the PCR fragments obtained in the first step (Fig. 1) was employed to remove the original sequence. On the other hand, herein template DNA was digested by restriction enzyme and purified from agarose gel. This treatment reduces the amplification of original sequence greatly. In addition, the use of single primer (T1) in the step 3 guarantees that newly synthesized fragments would be amplified exclusively. These modifications made possible to reduce the number of gel purification of PCR fragments, which was relatively time-consuming step, and to construct more than 20 mutant *gfp* genes in one day by a single researcher.

In addition, the reaction conditions of the *in vitro* transcription/translation were improved; for example, amino acid concentration was increased from 0.32 mM each to 0.7 mM each, which increased the amount of

Table 1. List of primers

S65OL 5'-GAAAGTAGTGACAAGTGTTGGC-3'
E22OL 5'-AAGAAGGACCATGTGGTCTCTCTTTTC-3'
F1T1 5'-GCGTACTAGCGTACCACGTGTCGACTATCTC GATCCCGCGAAATTAATACG-3'
R1T1 5'-GCGTACTAGCGTACCACGTGTCGACTTCCGG ATATAGTTCCTCCTTTCAG-3'
T1 5'-GCGTACTAGCGTACCACGTGTCGACT-3'
E222X primers
5'-GAAAAGAGAGACCACATGGTCCTTCTTNNNTTTG TAACAGCTGCTGGG-3'

Primer	Sequence	Primer	Sequence	Primer	Sequence	Primer	Sequence
E222A	GCG	E222C	TGC	E222D	GAT	E222E	GAG
E222F	TTT	E222G	GGC	E222H	CAT	E222I	ATT
E222K	AAA	E222L	CTG	E222M	ATG	E222N	AAC
E222P	CCG	E222Q	CAG	E222R	CGT	E222S	AGC
E222T	ACC	E222V	GTT	E222W	TGG	E222Y	TAT

S65X primers
5'-GCCAACACTTGTCACTACTTTCNNNTATGGTGTT CAATGCTTTTC-3'

Primer	Sequence	Primer	Sequence	Primer	Sequence	Primer	Sequence
S65A	GCG	S65C	TGC	S65D	GAT	S65E	GAG
S65F	TTT	S65G	GGC	S65H	CAT	S65I	ATT
S65K	AAA	S65L	CTG	S65M	ATG	S65N	AAC
S65P	CCG	S65Q	CAG	S65R	CGT	S65S	AGC
S65T	ACC	S65V	GTT	S65W	TGG	S65Y	TAT

produced protein *in vitro* (data not shown). Since PCR product can be used directly for the *in vitro* coupled transcription/translation reaction, mutant proteins was able to be obtained instantly, allowing a rapid and comprehensive examination of combinatorial mutations in proteins.

Measurement of Fluorescence of In Vitro Synthesized GFP

S65 and E222 are hydrogen bonded in the wild type GFP. This interaction is considered important to the protonation of chromophore, neutral and anionic, which corresponds to excitation peaks of 395 nm and 475 nm, respectively, of the wild type GFP [3]. To get more information of the effect of the mutual relationship between these residues on the fluorescent property, the two sites were mutated comprehensively.

Totally 400 proteins synthesized *in vitro* were analyzed fluorometrically. By the direct measurement using the reaction mixtures, 11 mutants: S65 changed to A (S65A), S65T, S65C, S65G, double mutation of S65 changed to T and E222 changed to Q (S65T-E222Q), S65A-E222A,-E222G, -E222C, -E222H, -E222S, and -E222Q) were found fluorescent. Most of the mutants lost an excitation peak of around 395 nm except for S65G mutant. Since some mutants having weak fluorescent signal might be undetectable due to relatively high background, the series of mutants having replacements of E222 together with S65S, A, or T, all of which showed relatively high fluorescence with the wild type residue E222, were purified and their fluorescence were analyzed. When combined with S65T mutation, E222 A, C, D, G, H, N, Q, S, V, or W mutant was found fluorescent additionally. When combined with S65A mutation, all replacements of E222 to any amino acids were found fluorescent, even though some of them showed a very weak signal. These results are summarized in Fig. 2. Among them, several mutants that had relatively high fluorescent signal were further analyzed, and the excitation and emission peaks were determined as listed in Table 2. Typical two-dimensional fluorescence plots are shown in Fig. 3. From the results of these experiments,

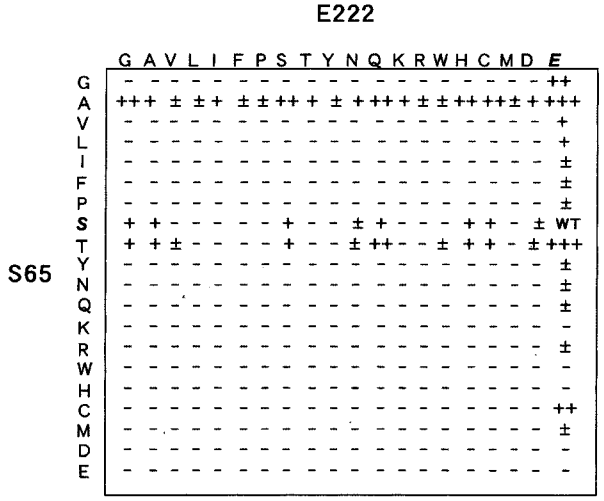

Fig. 2. Summary of GFP mutants. Relative fluorescent intensity against the wild type (the intensity of the wild type is expressed as 1) is indicated by +++, more than 1; ++, 0.5 to 1; +, 0.05 to 0.5; ±, less than 0.05, - undetectable.

Table 2. Emission and excitation peaks of GFP mutants

Mutants	Excitation	Emission
WT	397	507
	477	502
S65G	393	507
	485	506
S65A	476	507
S65V	480	506
S65L	485	506
S65C	481	506
S65T	490	510
S65A-E222G	480	502
S65A-E222A	481	502
S65A-E222C	480	503
S65A-E222H	477	500
S65A-E222S	477	501
S65A-E222Q	473	502
S65T-E222Q	483	505
E222G	480	500

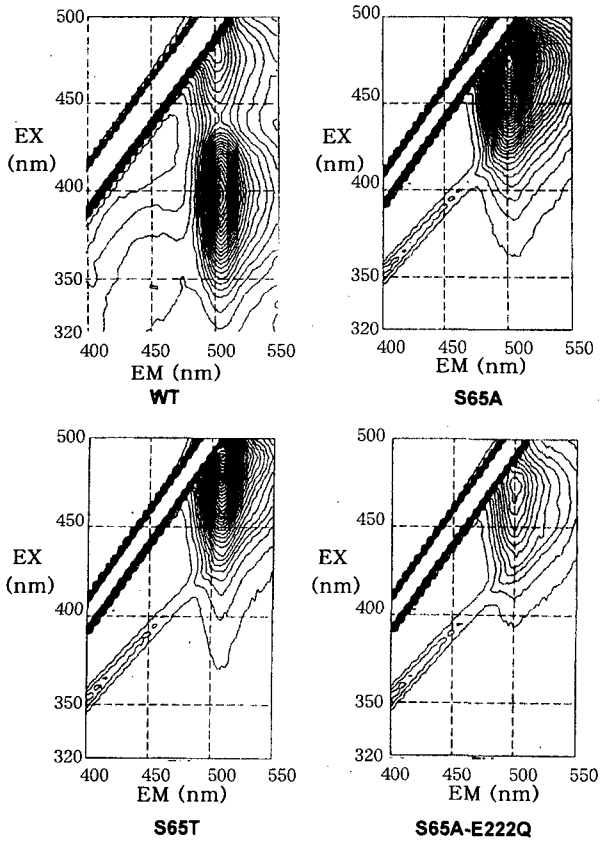

Fig. 3. Two-dimensional measurement of fluorescence of GFP mutants. Each were produced *in vitro* and purified as described in the text. Vertical and horizontal axes represent wave lengths of excitation and emission, respectively. Contour lines represent fluorescent intensity.

many mutant GFPs with fluorescence were found when combinatorial mutations were introduced at both 65th and 222nd position.

Common characters of the amino acid residues (S, G, A, V, L, T, C) that gave fluorescence to GFP at 65th position have small and non-charged side-chain, suggesting that spatial environment determines acceptability of each amino acid. In addition, S65A gave much higher fluorescence than S65G, which would suggest the importance of the methyl or the ethyl group of the side chain for the formation of the fluorophore.

Since S65 is involved in the fluorophore of GFP, it is likely that all mutations would have been already tried by other researchers. Actually if E222 was the wild type, mutants such as S65T and S65A were literally available. If S65 was changed to T, however, E222Q was also fluorescent. Interestingly, when S65 was replaced by A, the 222nd position was found to be tolerant to the replacements by all 20 amino acids, although the fluorescent intensity was strongly dependent on the character of amino acids: hydrophobic residues gave low signal but small and hydrophilic ones (G, A, S, Q, H, C) gave relatively high signal. All of these mutant proteins showed a similar emission maximum (around 505 nm) and a shifted excitation maximum at 480 nm. The high acceptability would be possibly explained by that the small side chain of A gives an enough space for a variety of side chains at 222nd position.

The hydrogen bond between E222 and S65 are supposed to be required to the excitation around 395 nm [3]. In our study, the all mutant except for S65G that potentially has no hydrogen bond lost the excitation but retained the excitation around 480 nm. Since the fluorescent intensity of S65G was relatively week, further analysis would be required to make any conclusion from this result.

By using the *in vitro* combinatorial mutagenesis, 399 mutants were comprehensively constructed by a sequential reaction only on a microplate exclusively without living cells in a high throughput manner. After purification of the first PCR fragment, only 3 to 4 h for PCR reaction and 1 h for transcription/translation were needed. As a result, a large number of mutants could be obtained within a day. Furthermore, the whole process can be automated since no laborious and complex step that needs human hands are included. The method presented here is a powerful tool to analyze the structure-function studies of proteins based on comprehensive mutations, which would give valuable information to design useful proteins.

Here *in vitro* coupled transcription/translation system was used instead of conventional *in vivo* expression systems. Although the expression level is not so high as the *in vivo* system, the obtained amount is enough for most of biochemical assay methods. More importantly, proteins that are difficult to produce as active form, such as a single-chain antibody [6], a Fab fragment [7], bacterial lipases [8,9] and phospholipase D [10], have been successfully synthesized avoiding the formation of aggregate by controlling oxidation conditions of the protein synthesis reaction of the *E. coli* S30 system. In addition, the *in vitro* expression system can be easily modified adjusting to target proteins. For example, addition of minor arginine and leucine tRNA increased production of proteins containing many such rare codons [11], and the use of protease-deficient mutant helped the stabilization of some proteins produced *in vitro* [12]. These reports clearly demonstrate the feasibility and usefulness of the *in vitro* expression system.

We have also developed another *in vitro* technology: the Single-Molecule PCR (SM-PCR), which enables a direct amplification from single DNA molecule [4,13, 14]. By direct combination of the SM-PCR and *in vitro* transcription/translation reaction, protein library can be constructed exclusively by the cell-free system only on micro-titer plates. In this case, introduction of random mutation is possible because each single DNA molecule is a start template for the PCR and the following protein synthesis reaction. This system is free from both transformation efficiency and the use of agar plates, which limit the library size of the conventional *in vivo* screening method. In addition, the obtained protein library is much more uniform than that obtained

by the *in vivo* methods [15].

The two systems we have developed are complementary, one is site-specific and targeted, but the other is rather compatible with random-mutation. They will provide, therefore, powerful tools with scientific and engineering fields of protein research by assisting comprehensive and high-throughput analysis.

CONCLUSION

We have demonstrated in this report that the *in vitro* combinatorial mutagenesis enables a rapid construction and analysis of all possible combinations of two specific sites of proteins, and provides comprehensive structure-function information easily. Combinatorial mutational analysis of 65th and 222nd positions in GFP using this methodology revealed that all possible amino acids were acceptable in place of E222 together with S65A, but not with S65G, suggesting the importance of the combination between these two sites. The results obtained here highlight the importance of comprehensive mutational analysis of function center of proteins and the usefulness of the *in vitro* system.

Acknowledgements This work was financially supported in part by a Grant-in-Aid for Scientific Research of the Japan Society of Promotion of Science (No. 12450332). The all DNA primers used in this study was kindly supplied by Nippn Flour Mils Co. Ltd.

REFERENCES

[1] Burks, E. A., G. Chen, G. Georgiou, and B. L. Iverson (1997) *In vitro* scanning saturation mutagenesis of an antibody binding pocket. *Proc. Natl. Acad. Sci. USA* 94: 412-417.

[2] Yang, J., Y. Koga, H. Nakano, and T. Yamane (2002) Modifying the chain-length selectivity of the lipase from *Burkholderia cepacia* KWI-56 through in vitro combinatorial mutagenesis in the substrate binding site. *Protein Eng.* 15: 147-152.

[3] Tsien, R. Y. (1998) Green fluorescent protein. *Ann. Rev. Biochem.* 67: 509-544.

[4] Ohuchi, S., H. Nakano, and T. Yamane, (1998) *In vitro* method for the generation of protein libraries using PCR amplification of a single DNA molecule and coupled transcription/translation. *Nucleic Acids Res.* 26: 4339-4346.

[5] Nakano, H., T. Shinbata, R. Okumura, S. Sekiguchi, M. Fujishiro, and T. Yamane (1999) Efficient coupled transcription/translation for PCR template by a hollow fiber membrane bioreactor. *Biotechnol. Bioeng.* 64: 194-199.

[6] Ryabova, L. A., D. Desplancq, A. S. Spirin, and A. Plückthun (1997) Functional antibody production using cell-free translation: Effects of protein disulfide isomerase and chaperones. *Nat. Biotechnol.* 15: 79-84.

[7] Jiang, X., Y. Ookubo, I. Fujii, H. Nakano, and T. Yamane (2002) Expression of Fab fragment of catalytic antibody 6D9 in an *Escherichia coli in vitro* coupled transcription/translation system. *FEBS Lett.* 514: 290-294.

[8] Yang, J., K. Kobayashi, H. Nakano, J. Tanaka, T. Nihira, Y. Yamada, and T. Yamane (1999) Modulator-mediated synthesis of active lipase of *Pseudomonas* sp. 109 by *Escherichia coli* cell-free coupled transcription/translation system. *J. Biosci. Bioeng.* 88: 605-609.

[9] Yang, J., K. Kobayashi, Y. Iwasaki, H Nakano and T. Yamane (2000) *In vitro* analysis of roles of a disulfide bridge and a calcium binding site in activation of *Pseudomonas* sp. strain KWI-56 lipase. *J. Bacteriol.* 182: 295-302.

[10] Iwasaki, Y., T. Nishiyama, Y. Kawarasaki, H. Nakano, and T. Yamane (2000) Importance of disulfide bridge formation on folding of phospholipase D from *Streptomyces antibioticus*. *J. Biosci. Bioeng.* 89: 506-508.

[11] Jiang, X., H. Nakano, T. Kigawa, T. Yabuki, S. Yokoyama, D. S.Clark, and T. Yamane (2001) Dosage effect of minor arginyl- and isoleucyl-tRNA on protein synthesis in an *E.coli in vitro* coupled transcription/translation system. *J. Biosci. Bioeng.* 91: 53-57.

[12] Jiang, X., K. Oohira, Y. Iwasaki, H. Nakano, S. Ichihara, and T. Yamane (2002) Reduction of protein degradation by use of protease-deficient mutants in cell-free protein synthesis system of *Escherichia coli*. *J. Biosci. Bioeng.* 93: 136-144.

[13] Nakano, H., K. Kobayashi, S. Ohuchi, S. Sekiguchi, and T. Yamane (2000) Single-step single-molecule PCR of DNA with a homo-priming sequence using a single primer and hot-startable DNA polymerase. *J. Biosci. Bioeng.* 90: 456-458.

[14] Koga, Y., K. Kobayashi, J. Yang, H. Nakano, and T. Yamane (2002) *In vitro* construction and screening of a *Burkholderia cepacia* lipase library using single-molecule PCR and cell-free protein synthesis. *J. Biosci. Bioeng.* 94: 84-86.

[15] Rungpragayphan, S., Y. Kawarasaki, T. Imaeda, K. Kohda, H. Nakano, and T. Yamane (2002) High-throughput, cloning-independent protein library construction by combining single-molecule DNA amplification with *in vitro* expression. *J. Mol. Biol.* 318: 395-405.

[Received March 15, 2002; accepted July 25, 2002]