Journal of Molecular Catalysis B: Enzymatic 63 (2010) 164–169

Contents lists available at ScienceDirect

Journal of Molecular Catalysis B: Enzymatic

journal homepage: www.elsevier.com/locate/molcatb

Codon optimization through a two-step gene synthesis leads to a high-level expression of *Aspergillus niger lip*2 gene in *Pichia pastoris*

Jiangke Yang [a,b,*], Liying Liu [b]

[a] *Key Laboratory of Molecular Biophysics of Ministry of Education, Huazhong University of Science and Technology, Wuhan 430074, China*
[b] *College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China*

ARTICLE INFO

Article history:
Received 15 August 2009
Received in revised form 7 January 2010
Accepted 8 January 2010
Available online 18 January 2010

Keywords:
Aspergillus niger
*lip*2
Two-step gene synthesis
Pichia pastoris
Expression

ABSTRACT

Aspergillus niger lipases are important biocatalysts for a broad range of industrial applications. To enhance the expression level of a newly cloned lipase gene *lip*2 of *A. niger* in *Pichia pastoris*, we applied codon optimization and synthesized the full length codon-optimized gene by a two-step gene synthesis strategy. This strategy consists of an assembly PCR for several small DNA fragments and enzymatic digestion and ligation steps to ligate these fragments into the full-length gene. First, the full-length *lip*2 gene was divided into three fragments F1 (237 bp), F2 (238 bp) and F3 (422 bp) with the additions of proper restriction sites, and separately amplified by assembly PCR reactions. Second, three PCR amplified fragments were digested and ligated into the full-length *lip*2 gene. In the two-step gene synthesis, synthesis of smaller DNA fragments resulted in a significant lower level of nonspecific mismatching among oligonucleotides and a very low mutational rate of the PCR products, demonstrating the superiority of the method. When compared with the originally cloned *lip*2 gene of *A. niger*, the new codon optimized *lip*2 gene expressed at a significantly higher level in yeasts after methanol induction for 72 h, and both the enzyme activity and protein content reached maximal levels of 191 U/ml and 154 mg/1, with 11.6- and 5.3-fold increases, respectively.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Lipases from *Aspergillus niger* possess various catalytic properties such as substrate specificity, enantioselectivity, and thermostability, making them important biocatalysts in a broad range of industrial applications [1–3]. As the natural extracellular products of *A. niger*, they are safe by the standard of the FDA, USA. Their activities require an optimal acidic pH, which is distinct from the pH optima of above 7 for most other bacterial and fungal lipases. Therefore, these unique properties make *A. niger* lipases the favorable selections in food processing [4].

Up to date, although several lipases have been purified and characterized from *A. niger* extracellular enzymes [5–7], most commercial products of *A. niger* lipases are still not pure preparations. The batch to batch variations in enzyme quality and quantity make it difficult to optimize the reaction conditions, to control and characterize the application processes [8]. Therefore, pure lipase preparations are highly desirable. In recent years, molecular cloning has greatly facilitated the large scale production and purification of many enzymes from extracellular components.

The first two *A. niger* lipase genes *lip*1 and *lip*2 have been cloned by our group [9]. Our initial enzymatic characterization revealed that these enzymes are potential biocatalysts for food processing, chiral resolution, and preparation of pharmaceuticals. Although we were able to successfully express these novel enzymes in *Escherichia coli* and *P. pastoris* [10], codon optimization is necessary to overcome the bias on codon usages between different microorganisms and achieve a high-level expression.

Currently, several methods of whole gene synthesis are generally applied to construct the gene with new codons. Among those [11–13], the more recently developed oligonucleotides assembly method has gained its popularity. First described by Stemmer et al. [13], a series of long oligonucleotides with overlapping sequences covering the complete sequence of both strands of a gene are synthesized, and then a full-length molecule is generated progressively by a single assembly PCR or a ligase chain reaction (LCR) [14]. To facilitate the process for oligonucleotide design and assemble, softwares such as Gene2Oligo [15], DNA2.0 (http://www.dna20.com), DNAWorks [16] and GeMS [17] have been developed. However, such one-step assembly PCR method has its limitations dealing with the complexity of the DNA secondary structures and the nonspecific mismatches among many oligonucleotides in one reaction.

* Corresponding author at: Huazhong University of Science and Technology, College of Life Science and Technology, 1037 Luoyu Road, Wuhan, Hubei 430074, China. Tel.: +86 27 87792214; fax: +86 27 87792213.
E-mail addresses: jiangke_yang@hotmail.com, jiangke.yang@gmail.com (J. Yang).

1381-1177/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.molcatb.2010.01.011

The oligonucleotides with shorter overlapped regions often cause more severe nonspecific annealing and result in errors of the full-length product such as truncated sequences caused by premature termination of the PCR reaction, and internal deletions or point mutations of the nucleotide sequences. The frequency of errors increases with the increase of gene length and the complexity of the secondary structure [18–21].

In order to overcome these problems, several strategies have been developed [18–21], but they have not been routinely used in gene synthesis. We now describe an improved two-step gene synthesis technique, in which several smaller DNA fragments are first synthesized by several separate assembly PCR reactions, and then are assembled into a full-length gene by restriction enzymatic digestion and subsequent ligation. Using this two-step method, we synthesized a codon optimized *A. niger lip*2 gene for the yeast expression and observed significantly elevated levels of both protein expression and enzyme activity in *P. pastoris* as compared with those from the gene with the original *A. niger* codons.

2. Experimental

2.1. Codon optimization, design and synthesis of oligonucleotides

Based on the native amino acid sequence of lipase gene *lip*2 (GenBank: FJ536288) of *A. niger* strain CICC 4009, which was deposited in China Centre of Industrial Culture Collection (CICC), the codons of the *lip*2 gene were optimized by replacing the codons predicted less frequently used in *Pichia* with the frequently used ones by DNA2.0 software (http://www.dna20.com).

The full-length sequence of *lip*2 gene was divided into three fragments F1 (237 bp), F2 (238 bp) and F3 (422 bp). Cla I restriction site was introduced at position 237 between F1 and F2, and Pst I was added at 475 site between F2 and F3. Oligonucleotide primers of 20–50 bp for assembly PCR of F1, F2 and F3 fragments were designed by Gene2Oliga software [15], and synthesis by Sangon Ltd., China (Supplementary data, Table S1).

2.2. Full-length gene assemble

A two-step method combining assembly PCR and enzymatic digestion/ligation process was used to synthesize full-length *lip*2 gene. In the first step, assembly PCR was conducted to separately assemble the oligonucleotides into fragments F1, F2 and F3. Assembly PCR reactions were carried out in a 50 μl volume containing 200 mM of each dNTP, 0.1 mM of each oligonucleotide, 1.5 mM $MgCl_2$, and 1 U of Pfu DNA polymerase (Promega). The PCR thermal cycling was set as a denaturation step at 94 °C for 2 min, and 30 cycles of 94 °C for 30 s, 52 °C for 30 s and 72 °C for 1 min, followed by a single incubation at 72 °C for 6 min. The products of assembly PCR were re-amplified by another round of PCR using two outer oligonucleotide primers in a 50 μl reaction containing 3 μl of assembly PCR mixture, 200 mM of each dNTP, 1 mM of each primer, 1 U of Pfu DNA polymerase in a buffer containing 1.25 mM of $MgCl_2$. The PCR products were then subjected to dA tailing and cloned into pMD18-T vector (Takara). Six positive clones were selected and sequenced to confirm their correct sequences.

PCR products of F1, F2 and F3 fragments were separately digested by respective enzymes, F1 by Cla I, F2 by Cla I and Pst I, F3 by Pst I, and then the digested fragments were ligated at 16 °C overnight. A 10 μl ligation mixture contained 10 U T4 DNA ligase and 20 ng each of the DNA fragment. Full-length *lip*2 gene was amplified by PCR with primer pair P1 (5′-CTTGAATTCGCTCCAGCTCCAGCACCTATGCAAA-3′, EcoR I site) and P2 (5′-CACATTAATTCGAAGAGGG, Not I site). A 50 μl reaction contained 200 mM dNTP, 0.1 mM P1 and P2 primers, and 1 U Pfu. The PCR condition was set as a denaturation step at 94 °C for 2 min, and 30 cycles of 94 °C 30 s, 52 °C 30 s, and 72 °C 1 min, followed by an extension step at 72 °C for 5 min.

2.3. Plasmid construction, transformation and recombinants selection

Vector pPIC9K (Invitrogen) was used for cloning and expression in *P. pastoris*. The plasmid contained an alcohol oxidase 1 promoter from *P. pastoris* fused to the α-mating factor from *Saccharomyces cerevisiae* for directing the protein to the secretary pathway. To make the *lip*2 fusion expression with α-factor, PCR product of *lip*2 was digested with EcoR I and Not I, and then inserted into pPIC9K to obtain plasmid pPIC9K-LIP.

Enzyme Sac I was used to linearize plasmid pPIC-LIP for the single crossover with *P. pastoris* genome to generate the methanol-utilized phenotype (Mut^+). About 6 μg of linearized DNA was mixed with 80 μl of competent cells, and the electroporation was conducted on Gene Pulser (Bio-rad) according to the manufacturer's suggestion for *S. cerevisiae*. Positive clones were initially selected by MD medium (1.34% yeast nitrogen base, 4×10^{-5}% biotin, 2% dextrose) plates and then checked by colony PCR. The insertion copy number of transformants was determined by G418 according to the description of Scorer et al. [22], and transformants with same copy number were selected.

2.4. Fermentation and lipase purification

A single colony of recombinant was picked and inoculated into 50 ml BMGY (1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer with pH 6.0, 1.34% yeast nitrogen base, 4×10^{-5}% biotin, 1% glycerol) medium, and grew at 28–30 °C in a shaking incubator (250–300 rpm) until the culture reached an OD_{600} of 3.0. The cells were harvested, and a portion was transferred into 50 ml BMMY (1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer pH 6.0, 1.34% yeast nitrogen base, 4×10^{-5}% biotin, and 0.5% methanol) medium to obtain a cell suspension with OD_{600} = 1.0. The cells were grown for another 5 d and expression of lipase was induced by methanol at a final concentration of 0.5% added every 36 h. Lipase activity was checked at all the time intervals. Purification of the lipase was conducted according to the description of Yang et al. [23].

2.5. Lipase activity and protein content assays

Lipase activity was quantified at pH 7.5 by free fatty acid titration with 50 mM NaOH after incubated in a thermostated vessel for 10 min. The assay mixture consisted of 5 ml 50 mM Tris–HCl buffer, 50 mM NaCl, 4 ml emulsified olive oil and 1 ml enzyme solution. One unit (U) of the activity was defined as the amount of enzyme liberating 1 μmol of fatty acid per min at 45 °C. Protein content of the fermentation broth was determined by the Bradford method [21].

3. Results

3.1. Gene synthesis design

Codon optimization by using frequently used codons in the host is an efficient measure to improve the expression level of heterologous gene [25,26]. In order to achieve a high-level expression of *A. niger lip*2 gene in *P. pastoris*, the codons of *A. niger lip*2 gene were replaced with those more frequently used codons by *P. pastoris* (Table 1 and Fig. 1). Two restriction sites Cla I (237 site) and Pst I (475 site) were introduced into the gene to divide the full-length DNA

Table 1
Codon usage of amino acid of original *lip*2 gene and the optimized gene in *Pichia pastoris*.

Amino acid	Codon	Frequency in *Pichia* (/1000)	Frequency in original gene (/1000)	Frequency in optimized gene (/1000)
Alanine, A	GCT	29.9	28.6	77.2
Cysteine, C	TGT	8.3	7.1	20.1
	TGC	4.4	14.3	0.0
Glycine, G	GGT	27.1	10.7	47.0
	GGC	8.5	42.9	0.0
Histidine, H	CAT	10.4	3.6	23.5
	CAC	9.1	17.9	0.0
Isoleucine, I	ATT	31.0	10.7	30.2
Leucine, L	TTG	31.9	10.7	40.3
	CTC	7.6	21.4	0.0
Proline, P	CCA	17.5	0.0	26.8
	CCG	3.9	21.4	0.0
Arginine, R	AGA	19.9	3.6	30.2
	CGT	6.7	7.1	0.0
	CGG	1.8	7.1	0.0
	CGC	2.3	7.1	0.0
Serine, S	TCT	23.9	7.1	60.4
	AGT	12.2	17.9	0.0
	AGC	7.3	32.1	0.0
	TCG	6.8	14.3	0.0
Threonine, T	ACT	23.7	14.3	57.0
	ACG	6.0	39.3	0.0
Valine, V	GTT	27.0	10.7	50.3

molecule into F1 (237 bps), F2 (238 bps) and F3 (422 bps) fragments. To keep the GC content of the synthetic gene within 40–60%, and the nucleotides A, T, G, and C dispersed evenly in the synthesized gene to eliminate AT- or GC-rich motifs, codons containing both AT and GC bases were selected when the differences between the codon frequencies were not significant. After codon optimization, the complexity of the RNA secondary of designed gene changed from the original 245 to 153, and the minimal free energy (MFE) increased from −363 to −304 kcal/mol (Fig. 2), as calculated by RNAfold software [27].

3.2. Two-step lip2 gene assemble

The steps of two-step *lip*2 gene synthesis are shown by the flowchart in Fig. 3. In the first step, assembly PCR was conducted to assemble the synthesized oligonucleotides covering both strands of

Fig. 1. Sequence comparison between the original (*lip*2) and the optimized (*lip*2-*syn*) genes. Dots in *lip*2 represent the same nucleotides between *lip*2 and *lip*2-*syn*. Underlined are the restriction site Cla I and Pst I.

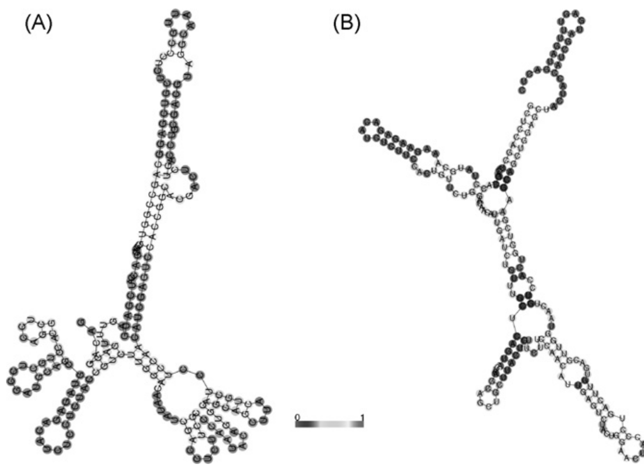

Fig. 2. Secondary structure of the first 200 bp of *lip*2 mature mRNA generated by the software RNAfolder. (A) Structure of original *lip*2 mRNA with the MFE is −67.6 kcal/mol and (B) structure of codon optimized *lip*2 mRNA with the MFE is −33.6 kcal/mol.

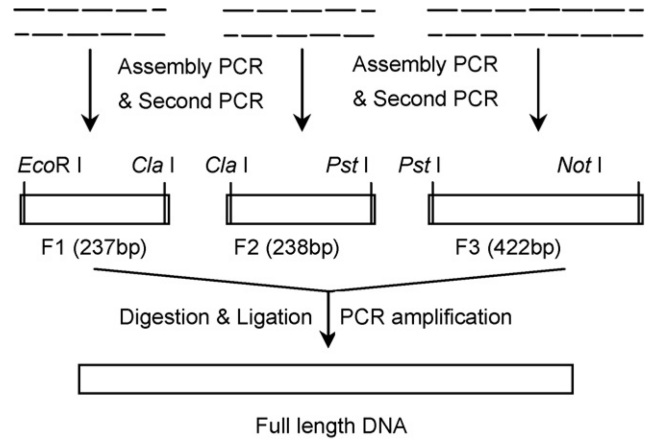

Fig. 3. Schema of the two-step *lip*2 gene synthesis method.

DNA molecule into the fragments F1, F2, and F3 (Fig. 4A). This step was similar to the general one-step assembly PCR gene synthesis method [13]. In the second step, an enzymatic digestion/ligation was conducted to ligate the three fragments into the full-length *lip*2 gene (Fig. 4B), as described in detail in Section 2.

3.3. Expression in P. pastoris

To evaluate the effect of the codon optimization, sample aliquots were obtained from the fermentation broths with recombinants carrying the new codon optimized gene or the original *A. niger lip*2 gene, and the enzymatic activities and the protein contents were determined and compared. On the tributyrin-MS plates, the lipases were expressed as secreting proteins from both the original and synthesized genes, and the lipase secretion from the recombinants carrying *lip*2-*syn* gene was much stronger than the cells with original gene (Fig. 5A). After methanol induction in BMMY medium in a flask scale, the contents of lipases in the fermentation broth were checked by SDS-PAGE (Fig. 5B) and the enzyme activities were measured and calculated. After inducible expression for 72 h, the lipase production curves (Fig. 6) showed that both the activity and the protein content in the supernatant of optimized *lip*2 recombinants reached the maximal levels of 191 U/ml and 154 mg/1, respectively,

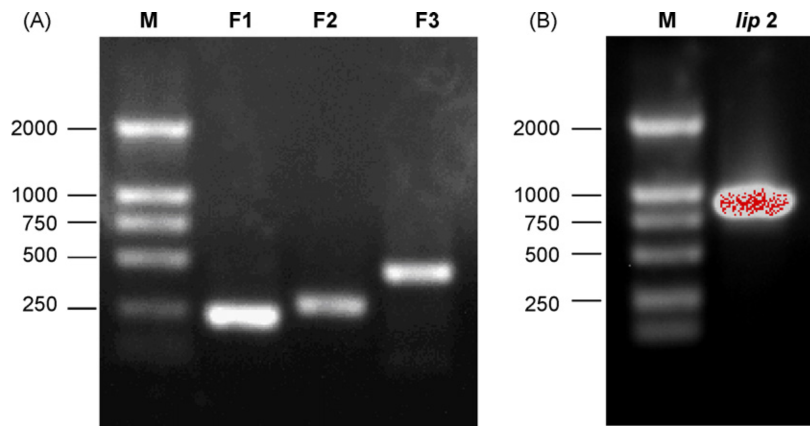

Fig. 4. Synthesis of long fragments and the full-length *lip*2 gene. (A) Fragments F1, F2 and F3 separately assembled by A-PCR and a second PCR. (B) PCR products of full-length *lip*2 gene amplified from the enzymatic digestion/ligation products.

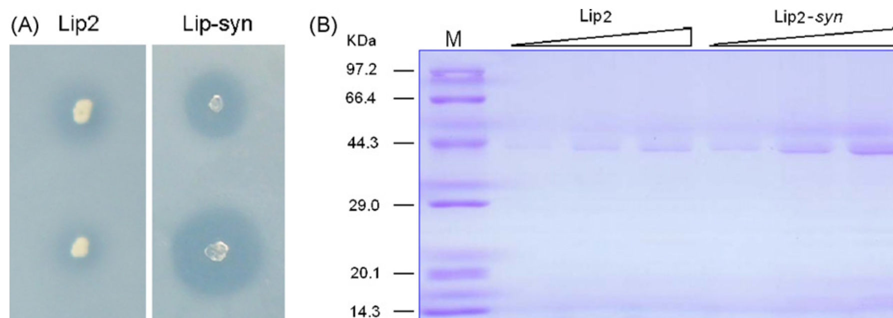

Fig. 5. Phenotype of the yeast recombinants carrying *lip*2 and *lip*2-*syn* genes and protein contents of the fermentation broth supernatant. (A) Phenotype of yeast recombinants on tributyrin-MS plates; (B) SDS-PAGE of protein in fermentation broth after methanol induced for 24, 48 and 72 h.

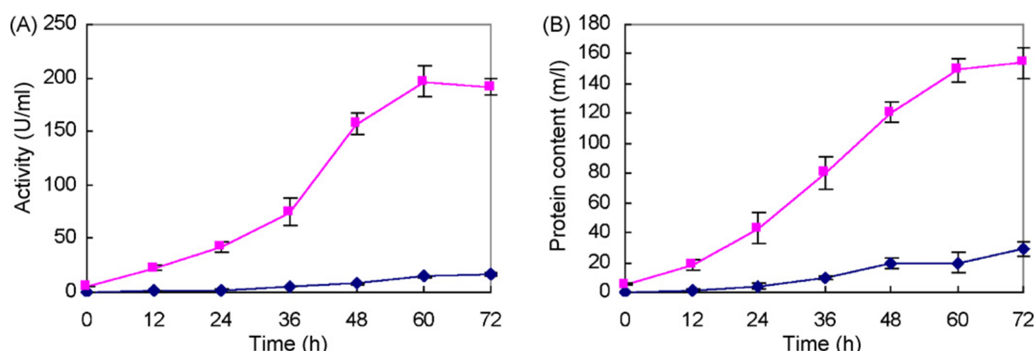

Fig. 6. Activity and protein content of the supernatant of fermentation broth from yeast recombinants carrying *lip*2 and *lip*2-*syn*. (A) Lipase activity of fermentation broth; (B) protein content of fermentation broth.

while the recombinants carrying original gene had only 16.5 U/ml and 29.3 mg/L, respectively.

4. Discussion

As an easy and simply system suitable for high density fermentation, *Pichia pastoris* is now broadly used to produce recombinant heterologous proteins, including a series of lipase genes from different organisms [8,23–25]. Due to the discrepancy of codon usages between the host and their original strains, the expression levels of these lipases hardly reach their optima. Codon usages between *A. niger* and *P. pastoris* are distinctly different. For examples, codons for amino acids Ala, Arg, Gly, Ser, Leu and Pro in *A. niger* are very infrequently used in *P. pastoris* (Table 1). Our codon optimization remarkably increased the expression level of the *lip*2 from the original 16.5 U/ml to a level of 191 U/ml in a small culture flask scale. Although our product yield is not as high as what Quyen et al. [28] produced from a larger scale 5 L bioreactor (309,000 U/L), it is still significantly better than most other reported results [29,30] under flask or even bioreactor fermentations.

The copy number of the gene cassette generally has the effect on the amount of protein expressed [31,32]. However, this effect was somehow rate-limiting and not so much effectively due to the complexity on the transcription, translation and secretion [33,34]. In this study, to make the recombinants of the original and codon optimized genes carrying the same copy number of gene cassette, the recombinants were screened by the same concentration of antibiotics G418 and then three recombinants sharing the uniform phenotypes were randomly picked out for the fermentation.

Long DNA sequences (>0.5 kb) are difficult to synthesize by the one-step assembly PCR due to the higher risk of mis-priming as the number of primers in the pool increases. Serious mismatches between the oligonucleotides can prematurely terminate the reaction and form premature DNA products [17–21]. To overcome these problems, a two-step gene synthesis method employing a dual asymmetric PCR (DA-PCR) or successive PCR to produce several fragments of a target gene, and then assembling them into a long DNA sequence by overlap-extension PCR (OE-PCR) has been developed and regarded as a main measure for long DNA sequence synthesis [18–21]. In this study, we adopted a two-step strategy combining assemble PCR and enzymatic digestion/ligation. Different from the other two-step methods described previously [18–21], in which fragments were assembled into a long DNA sequence by overlap-PCR (O-PCR), the additional enzymatic digestion/ligation step largely reduces the possibility of mutations caused by excessive PCR. In this study, all 18 clones (six clones for each fragment) showed no mutations when high fidelity DNA polymerase and Ultra-PAGE purified oligonucleotides were used in the system.

Codon optimization is an efficient measure to improve the expression level of other enzyme genes in heterologous host. In this study, the codons of *A. niger lip*2 gene was optimized by a two-step gene synthesis method. We showed that this is an efficient, simple and flexible method for synthesis of long DNA sequences. While the lipase production and enzyme activity from recombinants carrying optimized *lip*2 gene may be further improved under the batch-induced mode with a tighter controls of pH, methanol concentration and aeration during the production, the results presented here will greatly contribute to enhance production and characterization of biocatalytic enzymes; and our codon optimized *lip*2 gene will offer a greater value in various industrial applications.

Acknowledgements

We thank X. Yan and C. Ye for their technical help. This study was funded by a grant from Chinese High-tech Development Program (2007AA05Z417, 2009AA03Z232).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molcatb.2010.01.011.

References

[1] G. Carrea, A. Corcelli, G. Palmisano, S. Riva, Biotechnol. Bioeng. 52 (2000) 648–652.
[2] L. Sarıyar, D. Heperkan, Int. J. Food Sci. Technol. 38 (2003) 487–492.
[3] N. Saisubramanian, N. Edwinoliver, N. Nandakumar, N.R. Kamini, R. Puvanakrishnan, J. Ind. Microbiol. Biotechnol. 33 (2006) 669–676.
[4] E. Schuster, N. Dunn-Coleman, J. Frisvad, P. van Dijck, Appl. Microbiol. Biotechnol. 59 (2002) 426–435.
[5] J. Fukumoto, M. Iwai, Y. Tsujisaka, J. Gen. Appl. Microbiol. 9 (1963) 353–361.
[6] M.P. Tombs, G.G. Blake, Biochim. Biophys. Acta 700 (1982) 81–89.
[7] A. Sugihara, Y. Shimada, Y. Tominaga, Agric. Biol. Chem. 52 (1988) 1591–1592.
[8] L. Fernandez, I. Perez-Victoria, A. Zafra, P.L. Benitez, J.C. Morales, J. Velasco, J.L. Adrop, Protein Express. Purif. 49 (2006) 256–264.
[9] Z.Y. Shu, Y.J. Yan, J.K. Yang, L. Xu, Biotechnol. Lett. 29 (2007) 1875–1879.
[10] Z. Shu, M. Duan, J. Yang, L. Xu, Y. Yan, Biotechnol. Prog. 25 (2009) 409–416.
[11] R.C. Scarpulla, S. Narang, R. Wu, Anal. Biochem. 121 (1982) 356–365.
[12] W. Mandecki, T.J. Bolling, Gene 68 (1988) 101–107.
[13] W.P. Stemmer, A. Crameri, K.D. Ha, T.M. Brennan, H.L. Heyneker, Gene 164 (1995) 49–53.
[14] L.C. Au, F.Y. Yang, W.J. Yang, S.H. Lo, C.F. Kao, Biochem. Biophys. Res. Commun. 248 (1998) 200–203.
[15] J.M. Rouillard, W. Lee, G. Truan, X. Gao, X. Zhou, E. Gulari, Nucleic Acids Res. 32 (2004) W176–W180.
[16] D.M. Hoover, J. Lubkowski, Nucleic Acids Res. 30 (2002) e43.
[17] S. Jayaraj, R. Reid, D.V. Santi, Nucleic Acids Res. 33 (2005) 3011–3016.
[18] X. Gao, P. Yo, A. Keith, T.J. Ragan, T.K. Harris, Nucleic Acids Res. 31 (2003) e143.
[19] A.S. Xiong, Q.H. Yao, R.H. Peng, X. Li, H.Q. Fan, Z.M. Cheng, Y. Li, Nucleic Acids Res. 32 (2004) e98.
[20] L. Young, Q. Dong, Nucleic Acids Res. 32 (2004) e59.
[21] A.S. Xiong, Q.H. Yao, R.H. Peng, H. Duan, X. Li, H.Q. Fan, Z.M. Cheng, Y. Li, Nat. Protoc. 1 (2006) 791–797.

[22] C.A. Scorer, J.J. Clare, W.R. McCombie, M.A. Romanos, K. Sreekrishna, BioTechnology 12 (1994) 181–184.
[23] J. Yang, D. Guo, Y. Yan, J. Mol. Catal. B: Enzym. 45 (2007) 91–96.
[24] T. Zor, Z. Selinger, Anal. Biochem. 236 (1995) 302–308.
[25] G.C. Lee, L.C. Lee, V. Sava, J.F. Shaw, Biochem. J. 366 (2002) 603–611.
[26] S.W. Chang, G.C. Lee, J.F. Shaw, J. Agric. Food Chem. 54 (2006) 815–822.
[27] A.R. Gruber, R. Lorenz, S.H. Bernhart, R. Neuböck, I.L. Hofacker, Nucleic Acids Res. 36 (2008) W70–74.
[28] D.T. Quyen, C. Schmidt-Dannert, R.D. Schmid, Protein Express. Purif. 28 (2003) 102–110.
[29] S.O. Ramchuran, V.A. Vargas, R. Hatti-Kaul, Appl. Microbiol. Biotechnol. 71 (2006) 463–472.
[30] H. Yao, S. Yu, L. Zhang, K. Zuo, H. Ling, F. Zhang, K. Tang, Mol. Biotechnol. 38 (2008) 99–107.
[31] A. Vassileva, D.A. Chugh, S. Swaminathan, N. Khanna, Protein Express. Purif. 21 (2001) 71–80.
[32] R. Daly, M.T.W. Hearn, J. Mol. Recogn. 18 (2005) 119–138.
[33] H. Hohenblum, B. Gasser, M. Maurer, N. Borth, D. Mattanovich, Biotechnol. Bioeng. 85 (2004) 367–375.
[34] H. Marx, A. Mecklenbräuker, B. Gasser, M. Sauer, D. Mattanovich, FEMS Yeast Res. 9 (2009) 1260–1270.