World J Microbiol Biotechnol (2011) 27:1367–1372
DOI 10.1007/s11274-010-0587-0

ORIGINAL PAPER

MutS-mediated enrichment of mutated DNA produced by directed evolution in vitro

Tianying Zhong · Yafeng Zhou · Lijun Bi · Xian-En Zhang

Received: 29 March 2010 / Accepted: 20 September 2010 / Published online: 10 October 2010

Abstract Directed evolution in vitro is a powerful tool in the study and design of protein function. However, screening the desired mutants is a difficult task. To facilitate the screening, a method is proposed to eliminate wild type sequences and increase mutated DNA sequences, which is based on the preferential binding of MutS protein to heteroduplex DNA. Following error-prone PCR, amplified products are denatured and re-annealed to form heteroduplex and homoduplex DNA. Heteroduplexes are selectively bound to an engineered MutS protein and immobilized on a Strep-Tactin column. Homoduplexes are effectively removed by washing, and the final elution is enriched in mutated DNA sequences. One round of mutated DNA enrichment resulted in an about 2.3-fold of increase in mutation frequency compared to the control. The percentage of mutants rose from 44% in the control sample to 72% in the enrichment sample. Fluorescent assay by flow cytometry showed that the enrichment method increased the mutants with changed fluorescent activity by about 2.2-fold, which strongly justified the efficiency of enrichment in increasing mutants with functional changes. With reduced workload of screening and increased possibility of obtaining mutants with functional changes, the overall efficiency was improved by MutS-mediated enrichment of mutated DNA.

Keywords Directed evolution · MutS · Mutated DNA enrichment · Mutation frequency

T. Zhong · L. Bi (✉)
National Laboratory of Biomacromolecules,
Institute of Biophysics, Chinese Academy of Sciences,
100101 Beijing, China
e-mail: blj@sun5.ibp.ac.cn

T. Zhong
Jiangmen Polytechnic, 529090 Jiangmen, China

Y. Zhou · X.-E. Zhang (✉)
State Key Laboratory of Virology, Wuhan Institute of Virology,
Chinese Academy of Sciences, 430071 Wuhan, China
e-mail: x.zhang@wh.iov.cn

Introduction

In vitro directed evolution, or molecular evolution, is a powerful tool to generate proteins with improved properties or new functions that would otherwise take millions of years to naturally evolve (Bershtein and Tawfik 2008; Farinas et al. 2001). In this process, the "evolution" of DNA sequences is facilitated by the generation of random mutation (Neylon 2004). One of the most widely used methods in directed evolution is error-prone PCR (Lin-Goerke et al. 1997; Yu et al. 2009). Another core technique is known as DNA shuffling (Stemmer 1994; Xie and Wang 2008), which may be performed alone or in combination with error-prone PCR. A typical directed evolution procedure contains two major steps: (1) generation of a random diversity DNA library and (2) screening/selection. Step (1) is mature and universal, but Step (2) is the bottleneck and varies from case to case (Dower and Mattheakis 2002). The screening/selecting methods physically separate the individual bacteria or microorganisms present in the library, in which the individual is assayed for the desired property. Because the typical library size is many orders of magnitude larger than the number of variants, the screening step is the most labor-intensive and time-consuming. Though the screening/selection efficiency has been largely improved by the introduction of surface display technique and high throughput screening (HTS)

technology with the aid of robotics (Aharoni et al. 2005; Hertzberg and Pope 2000; Williams and Thorson 2008), there is still a large need for the continued development of efficient screening methods.

The workload of screening becomes more important when target proteins lack real-time profiling methods for screening due to the lacking of fluorescent or chromogenic activity. Moreover, in error-prone PCR, the mutation frequency is usually confined to around 0.5% (Wang et al. 2000), in which the majority of the DNA library are wild type sequences which cause much labor and time to obtain desired mutants. Therefore, a proper screening method is required to identify the desired mutations in a high background of wild type sequences. To this end, the workload may be reduced and the efficiency of screening step in directed evolution may be improved if the wild type sequences are eliminated and the mutated sequences are enriched. Different from the existing methods that enlarge the screening capacity with variant instrumental detection techniques, the mutated DNA enrichment method eliminates the wild type sequences which raise the noise of detection and cause unwanted work in the screening process.

In this report, a procedure is presented for reducing wild type sequences to enrich mutated DNA using immobilized MutS protein. MutS has a preference to bind mismatched or unpaired DNA, a property that has been widely used to detect gene mutation for medical diagnosis of diseases and SNPs of particular genes of drug resistances, human pathogenic and genetic diseases (Stanislawska-Sachadyn and Sachadyn 2005). Recently, methods have been developed to directly fish out subtle mutations in genomic DNA with MutS (Wang and Liu 2004; Gotoh et al. 2000). Using electrochemical impedance spectroscopy, the C-A mismatch was discriminated by MutS with the concentration of target strand as low as 100 pM (Gong et al. 2009).

Besides the extensive usage of MutS in mutation detection, this study first explores the mismatched DNA binding activity of MutS to improve the efficiency of directed evolution by removing the wild type sequences and enriching mutated sequences generated in the random diversity DNA library. With green fluorescent protein (GFP)-coding gene as a standard, the MutS-mediated enrichment results in a 2.3-fold of increase in mutation frequency compared to the control, and the percentage of mutant clones rises from 44 to 72%. Moreover, the flow cytometry data show that the enrichment increased the percentage of mutants with changed fluorescent activity by about 2.2-fold. With elimination of wild type sequences, the workload of screening is reduced and the possibility of obtaining mutants with changed functions is increased, thus the whole efficiency of directed evolution is improved.

Materials and methods

Construction and purification of engineered MutS fusion protein

The *mutS* gene was amplified by standard PCR and cloned into the *Bam*HI/*Sac*I sites of plasmid pQE30-linker peptide-Streptavidin-binding peptide to give plasmid pQE30-*mutS*-linker peptide-Streptavidin-binding peptide. The recombinant MutS fusion protein contains an N-terminal hexahistidine (His_6) tag and C-terminal Streptavidin-binding peptide (SBP) tag. *E. coli* strain M15 cells transformed with the plasmid were grown to OD_{600} −0.8 at 37°C, and fusion protein expression was induced with 0.8 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 5 h at 25°C. Proteins were purified using Ni–NTA column. The purity of MutS was determined by SDS–PAGE and the concentration was determined using the BCA method.

Construction of *gfp* library by error-prone PCR

Sequence of *gfp* (GenBank accession No. AY150807) was amplified by error-prone PCR reactions containing 10 ng template DNA, 400 nM each primer, 0.05 mM $MnCl_2$, and 2× *Taq* mix. PCR was carried out using 30 cycles of 94°C for 30 s, 58°C for 30 s and 72°C for 1 min followed by a final extension at 72°C for 5 min. Amplified *gfp* genes containing errors were induced to form heteroduplexes with wild type sequences by denaturing the purified PCR products at 95°C for 10 min followed by cooling to room temperature at a rate of 0.02°C/s. The re-hybridized DNA was fragmented to about 200 bp by being divided into two pools and digested to completion with HinfI plus HaeIII or NdeI plus KspAI. After purification, two pools of restriction enzyme-digested DNA were mixed to form the initial DNA substrates for the enrichment procedure.

Enrichment of mutated DNA by MutS

Sixty nano gram DNA/μl reacted with or without 50 μM MutS in binding buffer (20 mM Tris–HCl, pH 7.8, 150 mM NaCl, 5 mM $MgCl_2$). The mixtures were incubated at 4°C for 15 min to allow formation of MutS-heteroduplex complexes and then loaded onto a Strep-Tactin column (IBA) of equal volume pre-equilibrated in binding buffer. After 15 min incubation, the column was washed with 2 column volumes of binding buffer. Finally bound sample was eluted with 1 column volume of elution buffer (100 mM Tris–HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin). DNA in the final eluate was reassembled to full length of the *gfp* gene by PCR, and ligated to pUC19 vector. Ligation products were transformed into *E. coli* DH5α cells, then plated on solid LB media. Single

clones were randomly picked up for DNA sequencing. The control group was set by applying the same initial DNA sample to the enrichment process but without MutS.

Detection of fluorescence by flow cytometry

Flow cytometry experiments were carried out to detect the fluorescent activity of GFP. Cultures were grown to 0.4 OD_{600} and analysed by flow cytometry using a FACSCalibur system (BD Biosciences). Live cells were differentiated from dead cells and debris by analysis of the forward and side scattering properties of cells in the sample. The live cells were analysed for green fluorescence (530 nm). Cells transformed with plasmid pUC19 were used as the negative control and cells transformed with products of standard PCR of *gfp* ligated to pUC19 were the positive control.

Results

Scheme for MutS-mediated enrichment of mutated DNA

The procedure of enrichment of mutated DNA by MutS is shown in Fig. 1. DNA produced by error-prone PCR contains wild type sequences with less abundant mutant sequences. Following PCR, products are heated denatured and slowly cooled to room temperature to reform duplex DNA. Since most of the PCR products are wild type, the majority of duplexes formed will be complementary homoduplexes. However, the less abundant mutant DNA will anneal with wild type to form heteroduplexes containing mismatched base pairs. Heteroduplexes would be retained on the column by binding to immobilized MutS, when homoduplexes were removed during the washing step. Thus, the final elution from the column would be enriched with a higher concentration of mutated sequences.

Mutational efficiency of MutS-mediated enrichment of mutated DNA

Three parallel experiments of mutated DNA enrichment by MutS were carried out. Table 1 shows that one round of enrichment increased the mutation frequency about 2.3-fold compared to the control. The average percentage of mutants in total sequenced clones reached 72% after enrichment, much higher than the 44% in the control. The average mutated sites per mutant also increased from 2.7 to 3.7 by the enrichment. And the number of clones needed to be screened to identify one mutant decreased from 2.3 in control sample to 1.4 in enrichment sample.

Mutational spectra of sequenced mutants

To check whether the enrichment process would change the mutational spectra, the nature of the mutations in each samples was analysed in Table 2. The frequencies of

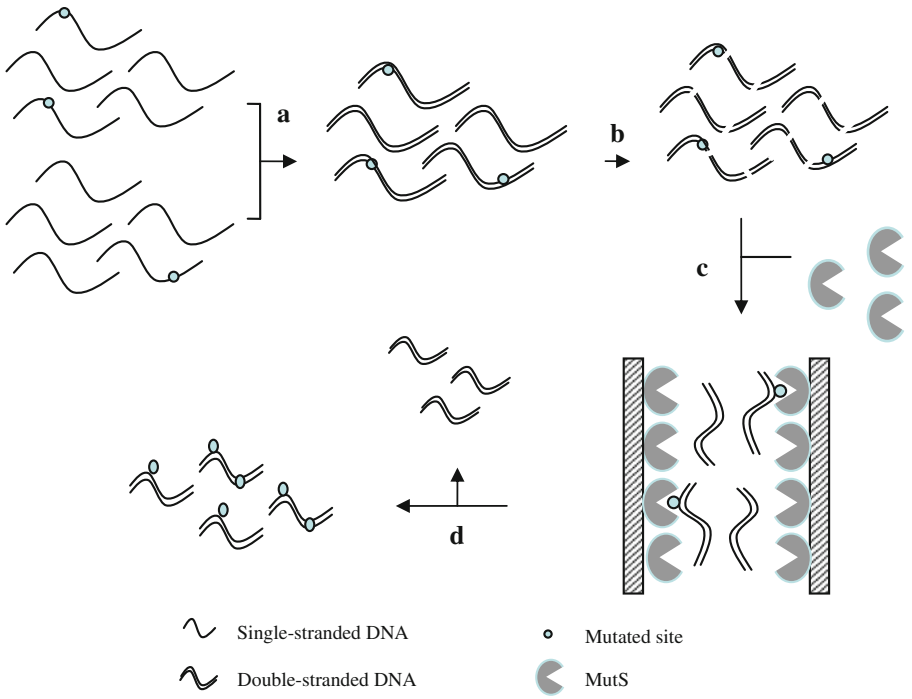

Fig. 1 The scheme of mutated DNA enriched by MutS. **a** The re-hybridization of error-prone PCR products results in mismatched heteroduplexes and complementary homoduplexes. **b** Fragmentation of DNA to length of about 200 bp. **c** The DNA is incubated with MutS and applied to the Strep-Tactin column. **d** Homoduplexes flow through the column with washing step, and heteroduplexes are eluted in the final elution, then subjected to the following steps for subcloning. Control is strictly set by applying the same initial DNA substrate to the process but without MutS

Table 1 Mutational efficiency of mutated DNA enrichment by MutS

		Sequenced clones	Mutant clones (%)	Screened clones per mutant	Mutated sites per mutant	Mutation frequency (10-3)	Efficiency of enrichment[a]
1	Control	27	52	1.9	2.3	1.6	2.1
	Enrichment	28	71	1.4	3.4	3.4	
2	Control	27	44	2.3	2.7	1.6	2.8
	Enrichment	26	81	1.2	4.0	4.5	
3	Control	35	37	2.7	3.1	1.6	2.0
	Enrichment	30	63	1.6	3.7	3.2	
Avg.[b]	Control	30	44	2.3	2.7	1.6	2.3
	Enrichment	28	72	1.4	3.7	3.7	

[a] The ratio of mutation frequency between Enrichment and Control

[b] The average of Group 1–3

Table 2 Comparison of base substitution spectra (presented in %)

		A→G/T→C	A→T/T→A	A→C/T→G	G→A/C→T	G→C/C→G	G→T/C→A	Deletion
1	Control	25	12.5	25	25	12.5	0	0
	Enrichment	64	0	0	24	6	0	6
2	Control	75	6	13	0	0	0	6
	Enrichment	64	5	7	22	0	2	0
3	Control	45	10	5	25	0	5	10
	Enrichment	71	6	3	20	0	0	0
Avg.[a]	Control	49	10	14	16	4	2	5
	Enrichment	66	4	4	22	2	1	2

No insertion was detected

[a] The average of Group 1–3

deletion, insertion and substitution mutations were calculated as the ratio between the corresponding events and the total mutations. Base substitutions were dominant when no insertion was detected in all mutants. AT to GC transition accounted for the major mutations in both control and enrichment groups, and the mutational bias showed no obvious variation after the enrichment process.

Changes in fluorescent activity of GFP

To analyse the activity of GFP, flow cytometry was used to assay the fluorescence of all living cells in each samples transformed with different DNA substrates. No stronger fluorescent GFP was detected in this experiment. Table 3 indicates that the percentage of non-fluorescent cells increased as the mutation frequency increased. Comparing to the standard PCR product of *gfp* (Group GFP, non-fluorescent cells, 8.0%), the net increase of non-fluorescent cells in Control was 8.6%, because of the mutations produced overwhelmingly by error-prone PCR. The net increase of non-fluorescent cells in Enrichment was 19.4%, where more mutants with changed activity occurred. Thus, comparing the net increase of non-fluorescent cells between Control and Enrichment group, the possibility of obtaining mutants with changes in fluorescent activity increased about 2.2-fold by the mutated DNA enrichment method.

Table 3 Flow cytometry analysis of cells transformed with different DNA

	Non-fluorescent cells (%)	Fluorescent cells (%)
pUC19[a]	100	0
GFP[b]	8.0	92.0
Control[c]	16.6	83.4
Enrichment[d]	27.4	72.6

[a] Plasmid pUC19

[b] Standard PCR product of *gfp* ligated into pUC19

[c] DNA from control group ligated to pUC19

[d] DNA after mutated enrichment ligated into pUC19

Discussion

Directed evolution in vitro is a powerful technique to study and engineer protein. But screening desired mutants from a large population is a difficult work, especially when a high percentage of wild type exists. Different from all methods used in directed evolution, this report first attempted to improve the screening step by enriching the mutated DNA sequences. The method was based on the selective binding of MutS to mismatched DNA. One round of mutated DNA enrichment resulted in an about 2.3-fold of increase in mutational efficiency. The percentage of mutants in total sequenced clones reached 72% of average after enrichment, much higher than the 44% of the control. With elimination of wild type sequences, this increased mutational efficiency did reduce the workload of subsequent screening procedure, therefore, improved the overall efficiency of directed evolution.

The fluorescent activity of the GFP libraries was detected by flow cytometry in a manner of high capacity. The difference in fluorescence between the *gfp* gene with and without fragmentation/reassembly step was not detected by flow cytometry in this study (data not shown), thus the changes in fluorescent activity were caused by the mutations in the *gfp* gene originated from error-prone PCR. Since no stronger fluorescent mutant was obtained, the changes in percentage of non-fluorescent cells could represent the changes in activity of GFP. When all living cells in each samples were analysed, the corresponding percentage of non-fluorescent cells increased about 2.2-fold in enrichment compared to the control. These data showed that the enrichment method increased the possibility of obtaining mutants with functional changes. In this way, the flow cytometry data strongly supported the mutational efficiency of the enrichment method, when the activity of protein was taken into account.

The specificity of MutS binding to a mismatched site, which decreased as the length of DNA substrates increased, was crucial to the enrichment method. When the full length of *gfp* gene (717 bp) was applied to the enrichment, the mutational efficiency resulted in a 1.6-fold increase compared to the control (data not shown). To improve the specific binding of MutS to mismatched DNA, a step of DNA fragmentation was added to the procedure, which not only raised the mutational efficiency to a 2.3-fold increase, but also made the enrichment process even more suitable for DNA shuffling in the way that the enriched mutated fragments could assemble more frequently, while wild type fragments were moved out by the washing step to form more mutants.

In an extreme case, the library of the *gfp* gene was produced by standard PCR using *Taq* polymerase with no manganese in the reaction, in which the mutation frequency was limited to a very low level and the wild type background was quite high. When this library was applied to the enrichment process, the mutation frequency raised from 0.16×10^{-3} to 0.22×10^{-3}, which was a 1.4-fold increase. This increase of mutational efficiency showed that the enrichment method was feasible to increase mutated DNA, even when a library with very small number of mutants was being screened.

There are at least two possible ways to improve the enrichment method. The amount of MutS in the reaction affected the binding pattern to the DNA significantly (Su et al. 2004). So varying the ratio of MutS to DNA might achieve better binding specificity of MutS to mismatched DNA. Alternatively, the transactivation model suggested that MutS scanned DNA for unpaired sequences and stalled on the mismatched point to trigger the subsequent repair steps (Junop et al. 2001). One way to improve the binding of MutS to the mismatched sequences, therefore, might be to introduce MutL, another mismatch repair protein that binds MutS to increase the stability of MutS-mismatch complex in an ATP-dependent manner (Joshi et al. 2000; Schofield et al. 2001).

Conclusion

In this study, the mutated DNA enrichment method resulted in an about 2.3-fold increase of mutation frequency compared to the control. And the fluorescent assay showed that the enrichment method increased the mutants with changed fluorescent activity by about 2.2-fold. These data strongly support that MutS-mediated mutated DNA enrichment was sufficient to increase the mutational efficiency and the possibility of mutants with functional changes. With this method, the whole efficiency of directed evolution was improved. Especially when the target proteins lacked an easy way for screening, the workload of screening becomes vital. In this case, the mutated DNA enrichment procedure provided a way to ease the workload of screening, even a twofold of increase might bring remarkable improvement.

Acknowledgments This work was supported by grants from the National Natural Science Foundation of China (No. 30670443) and the Chinese Academy of Sciences (No. KSCX1-YW-R-63, No. KSCX2-YW-G-017 and KZCX2-YW-420). The authors thank Chunchun Liu (the Protein Platform of Institute of Biophysics) for his assistance in flow cytometry analysis.

Conflict of interest The authors declare that they have no conflict of interest.

References

Aharoni A, Griffiths AD, Tawfik DS (2005) High-throughput screens and selections of enzyme-encoding genes. Curr Opin Chem Biol 9:210–216

Bershtein S, Tawfik DS (2008) Advances in laboratory evolution of enzymes. Curr Opin Chem Biol 12:151–158

Dower WJ, Mattheakis LC (2002) In vitro selection as a powerful tool for the applied evolution of proteins and peptides. Curr Opin Chem Biol 6:390–398

Farinas ET, Bulter T, Arnold FH (2001) Directed enzyme evolution. Curr Opin Biotechnol 12:545–551

Gong H, Zhong T, Gao L, Li X, Bi L, Kraatz HB (2009) Unlabeled hairpin DNA probe for electrochemical detection of single-nucleotide mismatches based on MutS-DNA interactions. Anal Chem 81(20):8639–8643

Gotoh K, Hata M, Miyajima M, Yokota H (2000) Genome-wide detection of unknown subtle mutations in bacteria by combination of MutS and RDA. Biochem Biophys Res Commun 268:535–540

Hertzberg RP, Pope AJ (2000) High-throughput screening: new technology for the 21st century. Curr Opin Chem Biol 4:445–451

Joshi A, Sen S, Rao BJ (2000) ATP-hydrolysis-dependent conformational switch modulates the stability of MutS-mismatch complexes. Nucleic Acids Res 28:853–861

Junop MS, Obmolova G, Rausch K, Hsieh P, Yang W (2001) Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. Mol Cell 7:1–12

Lin-Goerke JL, Robbins DJ, Burczak JD (1997) PCR-based random mutagenesis using manganese and reduced dNTP concentration. Biotechniques 23:409–412

Neylon C (2004) Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. Nucleic Acids Res 32:1448–1459

Schofield MJ, Nayak S, Scott TH, Du C, Hsieh P (2001) Interaction of *Escherichia coli* MutS and MutL at a DNA mismatch. J Biol Chem 276:28291–28299

Stanislawska-Sachadyn A, Sachadyn P (2005) MutS as a tool for mutation detection. Acta Biochim Pol 52:575–583

Stemmer WP (1994) Rapid evolution of a protein in vitro by DNA shuffling. Nature 370:389–391

Su X, Robelek R, Wu Y, Wang G, Knoll W (2004) Detection of point mutation and insertion mutations in DNA using a quartz crystal microbalance and MutS, a mismatch binding protein. Anal Chem 76:489–494

Wang J, Liu J (2004) Directed fishing out subtle mutations in genomic DNA with histidine-tagged *Thermus Thermophilus* MutS. Mutat Res 547:41–47

Wang D, Zhao C, Cheng R, Sun F (2000) Estimation of the mutation rate during error-prone polymerase chain reaction. J Comput Biol 7:143–158

Williams GJ, Thorson JS (2008) A high-throughput fluorescence-based glycosyltransferase screen and its application in directed evolution. Nat Protoc 3:357–362

Xie T, Wang Q (2008) Directed evolution of streptomyces lividans xylanase B toward enhanced thermal and alkaline pH stability. World J Microbiol Biotechnol 25:93–100

Yu H, Li J, Zhang D, Yang Y, Jiang W, Yang S (2009) Improving the thermostability of N-carbamyl-D-amino acid amidohydrolase by error-prone PCR. Appl Microbiol Biotechnol 82:279–285