Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biosensors and Bioelectronics 21 (2005) 135–144
BIOSENSORS & BIOELECTRONICS
www.elsevier.com/locate/bios

Construction and characterization of different MutS fusion proteins as recognition elements of DNA chip for detection of DNA mutations

Li-Jun Bi[a,b], Ya-Feng Zhou[b], Xian-En Zhang[a,b,*], Jiao-Yu Deng[b], Ji-Kai Wen[b], Zhi-Ping Zhang[b]

[a] *Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Chaoyang District, Beijing 100101, People's Republic of China*
[b] *Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, People's Republic of China*

Received 8 June 2004; received in revised form 17 August 2004; accepted 20 August 2004
Available online 13 October 2004

Abstract

Three MutS fusion systems were designed as the mutation recognition and signal elements of DNA chips for detection of DNA mutations. The expression vectors containing the encoding sequences of three recombinant proteins, Trx-His_6-GFP-$(Ser\text{-}Gly)_6$-MutS (THGLM), Trx-His_6-$(Ser\text{-}Gly)_6$-Strep tagII-$(Ser\text{-}Gly)_6$-MutS (THLSLM) and Trx-His_6-$(Ser\text{-}Gly)_6$-MutS (THLM), were constructed by gene slicing in vitro. THGLM, THLSLM and THLM were then expressed in *Escherichia coli* AD494(DE3), respectively. SDS-PAGE analysis revealed that each of the expected proteins was ~30% of the total bacterial proteins. The recombinant proteins were purified to the purity over 90% by immobilized metal (Co^{2+}) chelation affinity chromatography. Bioactivity assay indicated that three fusion proteins retained the mismatch-binding activity and the functions of other fusion partners. DNA chips arrayed both mismatched and unpaired DNA oligonucleotides as well as *rpoB* gene from *Mycobacterium tuberculosis* were prepared. THGLM, THLSLM and THLM that was labeled with Fluorolink™ Cy3 reactive dye, were then used as both mutation recognition and labeling elements of DNA chips. The resulting DNA chips were used to detect the mismatched and unpaired mutations in the synthesized oligonucleotides and single base mutation in *rpoB* gene of *M. tuberculosis* that is resistant to rifamycin.

Keywords: MutS fusion proteins; DNA chip; Mutation detection

1. Introduction

Most human genetic diseases, cancers and polymorphisms result from gene mutations including single-base substitutions and small insertions or deletions of bases in the genome (McCready et al., 2002; Beltran-Valero De Bernabe et al., 2002; Howard et al., 2002; Schaller et al., 2002). The rapid rate of discovery of disease genes increases the need for developing high-throughput mutation detection methods allowing screening of many individuals at multiple loci. Among a number of approaches for detection of DNA mutations, DNA direct sequencing is regarded as being most reliable, but quite time consuming and costly (Wong et al., 1987; Saiki et al., 1985). The other widely used methods for small genetic alteration detection include single-stranded conformation polymorphism (Sugano et al., 1995), denaturing gradient gel electrophoresis (Guldberg and Guttler, 1993), enzyme or chemical mismatch cleavage (Deeble et al., 1999; Hacia, 1999). These methods are all based on the use of PAGE and quite labor intensive, which significantly limits their suitability in automated applications requiring rapid screening of samples.

The MutS protein is a central element of DNA mismatches repair (MMR) system responsible for detection of pre-mutation changes in cells by DNA binding at the sites containing mismatched bases. MutS was reported to specifically recognize and bind all possible single-base mismatches as well as 1–4 base insertion/deletion loops with varying affinities in vitro independent of other proteins or cofactors (Lu et al., 1983; Lieb, 1987; Jiricny et al., 1988). Many meth-

* Corresponding author. Tel.: +86 10 58881508; fax: +86 10 64888464.
E-mail address: zhangxe@mail.most.gov.cn (X.-E. Zhang).

0956-5663/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.bios.2004.08.045

ods based on the MutS protein for detection of DNA mutations have been described. Solution phase approach using the specificity of the MutS protein mismatch protection from exonuclease (MutEx) (Ellis et al., 1994) for mutation detection was first reported to map allelic variants in *exon 11* of the cystic fibrosis transmembrane regulator gene, and then similar methods were developed soon afterwards (Lishanski et al., 1994; Smith and Modrich, 1996; Parsons and Heflich, 1997, 1998; Sachadyn et al., 2000; Beaulieu et al., 2001). The MutEx assay gives excellent signal to noise rations, but weak MutS binding to some mismatches may cause pseudo negative results. Wagner et al. demonstrated that MutS immobilized on nitrocellulose exhibited enhanced ability to discriminate between DNA with and without mismatches relative to MutS in solution and detected single nucleotide polymorphism (Wagner et al., 1995; Wagner and Dean, 2000). A fusion protein of MutS with a biotinylated peptide domain was employed by Geschwind et al. (1996) to successfully screen known homozygotes and heterozygotes for a single base-pair polymorphisms in *KCNA1* gene from human genomic DNA (Genebank accession L02750) using a non-gel based assay. The use of sensor methods to monitor MutS-DNA interactions was also reported in the recent years. Gotoh et al. (1997a,b) developed a MutS-based surface plasmon resonance optical biosensor for rapid and highly sensitive detection of point mutations. Han et al. (2002) also applied the MutS protein immobilized onto an Au electrode to detect GT mismatch and deletion mutation in the double-stranded DNA. Su et al. (2004) combined the MutS based mismatch recognition with quartz crystal microbalance (QCM) measurement for detection of point mutation and insertion mutations in DNA. Other interesting attempts include using representational difference analysis (RDA) techniques (Gotoh et al., 2000), atomic force microscopy (AFM) (Tanigawa et al., 2000; Sun and Yokota, 2000) to analysis or imaging the mutation. By using AFM technique the positions of the bound MutS proteins along the DNA molecules were determined by calculating the distance from one of the DNA termini. At this stage, these papers reflect a growing interest in DNA mismatch repairs as potential tools for screening genomes for polymorphisms and disease-causing mutations, although none of the MutS-based mutation detection methods have enjoyed widespread application for high-throughput, parallel and multiplex analysis. In recent years, DNA chip technology developed rapidly and has shown its great advantage in the parallel analysis of multiple sequences, so it is a point to use this feature in combination with MutS for detection of DNA mutations. Behrensdorf et al. (2002) recently reported a DNA chip for mutation detection with a fluorescence-labeled mismatch-binding protein as the recognition element, employing an electronically controlled DNA hybridization on open-format arrays. As an alternative, we proposed a protein chip format for the same purpose using a genetically modified MutS protein (Bi et al., 2003). Both DNA chip format and protein chip formats showed promising prospect in the parallel and multi analysis of mutations.

As a further approach, in this study we constructed three different MutS fusion proteins by gene splicing, which were Trx-His_6-GFP-$(Ser\text{-}Gly)_6$-MutS (THGLM), Trx-His_6-$(Ser\text{-}Gly)_6$-Strep tagII-$(Ser\text{-}Gly)_6$-MutS (THLSLM) and Trx-His_6-$(Ser\text{-}Gly)_6$-MutS (THLM), respectively. The backbone of the three fusions was built up with a package of Trx for solubilization, a His_6 for affinity purification, a $(Ser\text{-}Gly)_6$ for linkage and spacing, and MutS for recognizing and binding to the target DNA. To generate detectable signal, THGLM contained green fluorescent protein (GFP), THLSLM contained Strep tagII that could specifically interact with streptavidin labeled with alkaline phosphatase, and THLM was subsequently modified with Cy3 dye. Thus, all these fusion proteins possess two functions: recognizing and binding mutations and generating signals. The bioactivities of these various fusion configurations were characterized, and comparison study in using them to build the reliable DNA chips was carried out. The results are reported herein.

2. Materials and methods

2.1. Materials

Escherichia coli AD494 (DE3) and plasmid pET32a (+) were from Novagen Company. *gfp* gene is kindly donated by Dr. Li-Xin Ma. PCR primers used for cloning *gfp* and *mutS* gene were synthesized by Sangon Company (Shanghai, China). Ampicillin, Kanamycin sulphate, BSA, sopropylthio-β-D-galactoside (IPTG) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyronoside (X-gal) were purchased from Sigma Chemical Company. Restriction enzymes, DNA polymerase, and T4 DNA ligase were obtained from Takara and Promega Company (Shanghai, China). PCR purification mini kit and gel extraction mini kit were purchased from Huaxun Company. *E. coli* DH5α used for all bacterial transformations and plasmid propagations and *E. coli* K-12 are stored by our laboratory. ALON metal affinity resin was purchased from Novagen Company. Cy3 biofunctional dye kit was obtained from Amersham Pharmacia Biotech. Inc. All other reagents used were of analytical-reagent grade.

2.2. Construction of expression vectors of the MutS fusion proteins

All oligonucleotides were designed to incorporate proper restriction enzymatic sites for cloning. Primers G-1 (5′-AGATCTGATGAGTAAAGGAGAAGAACTTT-3′) and G-2 (5′-GATATCATCTTTGTATAGTTCATCCATG-3′), spanning the *gfp* gene, were synthesized to amplify *gfp* gene. Restriction enzymatic sites *Bgl*II and *Eco*RV were introduced to the N- and C-terminal primers. Primers M-1 (5′-GAATTCATGAGTGCAATAGAAAATTTCGAC-3′) and M-2 (5′-AAGCTTTATTTTTATTTGATTCGTCAGTTAT-3′), spanning the *mutS* gene, were used to amplify *mutS* gene. Restriction enzymatic sites *Eco*RI and *Hin*dIII were

introduced to the N- and C-terminal primers. PCR products were purified with PCR purification mini kits and then cloned into the pGEM-Ti vector separately to yield the plasmids pGEM-T-*gfp* and pGEM-T-*mutS*. The resulting plasmid pGEM-T-*mutS* was digested with *Eco*RI and *Hin*dIII, and MutS encoding fragment was recovered and inserted into the *Eco*RI/*Hin*dIII site of pET32a (+), yielding the plasmid pET32a-*mutS*. Oligonucleotides L1 (5′-ATC-AGCGGCTCAGGATCTGGATCAGGATCTGGCG-3′) and L2 (5′-AATTCGCCAGA TCCTGATCCAGATCCTAGCC-GCTGAT-3′) were annealed at 70 °C for 5 min to form a duplex DNA fragment containing a linker (-(Ser-Gly)$_6$-**)** encoding sequence and *Eco*RV and *Eco*RI sites at its N and C-termini, respectively. The plasmid pGEM-T-*gfp* was digested with *Bgl*II and *Eco*RV, and GFP encoding fragment was recovered. The encoding sequences of GFP and -(Ser-Gly)$_6$- were then inserted to replace the corresponding sequence between *Bgl*II and *Eco*RI in pET32a-*mutS*, and yielded the expression vector pET32a-*gfp*-(*Ser-Gly*)$_6$-*mutS*. The construction of expression vector of the fusion protein THLSLM has been described in our previous publication (Bi et al., 2003). In the case of expression vector of THLM, the -(Ser-Gly)$_6$- encoding fragment described above was directly inserted to replace the corresponding sequence between *Bgl*II and *Eco*RI in the recombinant plasmid pET32a-*mutS*, and yielded the expression vector pET32a-(*Ser-Gly*)$_6$-*mutS*.

2.3. Expression and purification of the MutS fusion proteins

Expression of the three fusion proteins was accomplished by inducing the positive transformants with IPTG in mid-exponential phase. The positive transformant cells were cultured in LB medium containing Ampicillin and Kanamycin sulphate for 4 h and continue to incubate with the addition of the final concentration IPTG of 1 mM for 5 h. The cells were harvested by centrifugation at 4000 × *g* for 30 min.

Purification of the fusion proteins was performed with His-Bind resin slurry. All steps were carried out at 4 °C. The preparation of Co^{2+}-charged chromatography resin column included washing with 10 column-volumes of sterile H_2O and equilibrating with 1 × extraction/washing buffer (50 mM sodium phosphate, 300 mM NaCl, pH 7.0).

The induced cells harvested by centrifugation were washed with deionized water and resuspended in 10 mL 1 × extraction buffer per 100 mL of cell culture, followed by sonication. The lysates were centrifugated at 12,000 × *g* for 30 min. Cell-extract supernatant was loaded onto the affinity chromatography resin column. The column was then washed with 10 column-volumes of 1 × extraction/wash buffer. The expected protein was eluted by washing the resin with 1 × elution buffer (50 mM sodium phosphate, 300 mM NaCl, 150 mM imidazole, pH 7.0). The elution buffer containing purified fusion proteins was dialyzed against 1 L of 1 × dialyzing buffer (50 mM Tris–HCl, pH 7.2, 100 mM KCl, 1 mM DTT and 1 mM EDTA) overnight. Concentration of the fusion proteins was determined using Bio-Rad assay kit with BSA as the standard, and purity of the proteins were evaluated by SDS-PAGE.

2.4. Labeling of the fusion protein THLM with Cy3 biofunctional dye kit

The labeling of the fusion protein THLM was performed for 30 min at room temperature in a 1 mL reaction mixture containing THLM and Cy3 biofunctional dye. THLM to be labeled was dissolved at concentration of 5 mg/mL in sodium carbonate–sodium bicarbonate buffer and this solution was added to the dye vial. The vial was then capped and mixed thoroughly. Care should be taken to prevent foaming of the fusion protein solution. The reaction mixture was incubated with additional mixing approximately every 10 min for 30 min at room temperature. Cy3-labeled THLM can be prepared from the excess, unconjugated dye by gel filtration chromatography. It is convenient to pre-equilibrate the column with phosphate-buffered saline and to elute the THLM protein using the same buffer. Two pink bands developed during elution. The faster moving band was Cy3-labeled THLM while the slower band was free dye. Cy3-labeled proteins were stored at −20 °C.

2.5. Mismatch-binding bioactivity assay of the MutS fusion proteins and Strep tagII bioactivity assay of the fusion protein THLSLM

The mismatch-binding activity of the fusion proteins and the Strep tagII bioactivity assay has been described in our previous publication (Bi et al., 2003).

2.6. Preparation of DNA chips

The glass slides were first cleaned with 5% K_2CrO_4 and 98% H_2SO_4, and then etched with 20% hydrofluoric acid to form desired arrays. The etched slides were washed with deionized water to remove the hydrofluoric acid. The cleaned slides were immersed into 25% ammonia solution overnight and then rinsed with millipore water for 10 min and briefly rinsed with anhydrous ethanol. The slides were then immersed in a mixture of 1% 3-mercaptopropyl trimethoxysilane (Sigma), 95% ethanol, and 16 mM acetic acid (pH 4.5) for 30–60 min. After rinsing with 95% ethanol and 16 mM acetic acid (pH 4.5) immediately, the silanized slides were cared under dry nitrogen at room temperature overnight.

The 5′-disulfide probes (5′-AATAGTTCTCAGGT-XGACGGATCTGGACAC-3′) were diluted with 500 mM $NaHCO_3/Na_2CO_3$ buffer (pH 9.0) and arrayed onto the mercaptosilane-coated wells. The slides were incubated in a humid chamber for over 14 h at 18 °C, followed by washing in TNTW buffer (10 mM Tris–HCl, pH 7.5,

Table 1
PCR primers

Oligonucleotide description	Sequences
130 bp *rpoB* gene fragment forward primer	5′-GGCGATCAAGGAGTTCTTC-3
130 bp *rpoB* gene fragment reverse primer	5′-GCACGCTCACGTGACAGACC-3
320 bp *rpoB* gene fragment forward primer	5′-GCGAGCTGATCCAAAACCA-3′
320 bp *rpoB* gene fragment reverse primer	5′-GGTTTCGATCGGGCACAT-3′
612 bp *rpoB* gene fragment forward primer	5′-TGTTGAAAACTTGTTCTTCA-3′
612 bp *rpoB* gene fragment reverse primer	5′-AGCCGATCAGACCGATGT-3′

150 mM NaCl, and 0.05% Tween 20) for 5–10 min. The hybridization mixture containing corresponding oligonucleotides was loaded on the chip and sealed in a hybridization chamber. The chamber was incubated at 37 °C for 3 h.

130-, 320-, and 612-bp-long *rpoB* gene were obtained by PCR, respectively, from wild-type and rifampin-resistant *Mycobacterium tuberculosis* using primers listed in Table 1. All the forward primers were modified with disulfide at their 5′ ends. PCR products were analyzed by electrophoresis in agarose gels. The denatured PCR products from wild-type and mutant were mixed and incubated at 70 °C to produce duplex *rpoB* gene fragments containing a mismatch.

After removal of hybridization buffer, the chip was blocked with 3% (w/v) BSA-0.5% (v/v) Tween 20 in 1 × assay buffer for 20 min, and then washed three times with 1 × assay buffer with 0.1% Tween 20 and was air dried. 0.5 mg/mL of the fusion protein, THGLM, Cy3-labeled THLM or THLSLM was added to the wells and incubated at 4 °C for 30 min to specially bind mismatched or unpaired duplex DNA. Following removal of the fusion protein not binding to duplex DNA, the fusion protein binding to duplex DNA was detected by corresponding method.

2.7. Signal measurement

A LEICA DMIRB microscope system (Leica Microsystems and the Eppendorf Company) was used to observe the green fluorescence signal of THGLM on the DNA chips. Images were collected using a Micropublisher color camera (Nikon Company) and analyzed with origin 6.0. The color signal images developed from enzymatic labeling of THLSLM on the DNA chip were collected using Olympus camera. The data were analyzed by origin 6.0. Cy3 fluorescence images from the Cy3-labeled THLM on the DNA chip were measured by GenePix 4000B (Axon Instrument) fluorescence Scanner with the sensitivity of 0.1 fluorophores/μm^2 and the data was analyzed with GenePix 4.0 analysis software (Axon Instruments).

3. Results

3.1. Expression and purification of THGLM, THLSLM and THLM

The expression vectors of THGLM, THLSLM and THLM were constructed by gene splicing in vitro (Fig. 1). The fusion proteins were over-expressed by inducing with IPTG under the control of a T7-phage promoter. The total proteins of the cell were analyzed by SDS-PAGE (Fig. 2). The protein bands corresponding to THGLM, THLSLM and THLM were visualized at 135 kD, 112 kD and 110 kD, respectively. To optimize the production conditions, effects of growth temperature, inducing time and IPTG concentration on the yields of the fusion proteins were investigated. The maximum yields of the fusion proteins, THLSLM and THLM, were obtained when the culture was induced with 1.0 mM IPTG and then incubated to grow at 30 °C for 5 h, but the maximum yield of the fusion protein THGLM was obtained when the culture was incubated to grow at 28 °C (data not shown). Purification of the fusion proteins was performed using His-Bind resin slurry according to the instruction of manufacturer. The elution fractions containing the desired proteins were collected and then dialyzed to remove salt overnight.

3.2. Detection of DNA mutation with THGLM-based DNA chip

The procedure of the detection was schematically shown in Fig. 3. (I) The 5′-disulfide DNA probes were immobilized on the SH-modified glass support; (II) corresponding oligonucleotides were added to hybridize with the immobilized DNA probes to form duplex DNA, which might be mismatched or complementary DNA; (III) THGLM was then added to recognize and bind to the mismatched duplex DNA. It was then able to discriminate between mismatched DNA and complementary DNA by measuring the difference of the intensity of green fluorescence signal generated from GFP with fluorescence microscope. The resulting DNA chips were used to detect the mismatched and unpaired mutations in the synthesized oligonucleotides, as well as the single base mutation in *rpoB* gene from *M. tuberculosis*. The results showed that THGLM-based DNA chip could readily detect nearly all the mismatches tested here (Fig. 4A). The pixel intensity of fluorescence of the fragments containing a GT GG, AA, and GA mismatches were 46,119, 45,225, 39,252 and 44,291, respectively, while those of the control experiments were only 2221 and 4808, respectively.

The binding affinities of the fusion protein THGLM to oligonucleotides containing 1–4 unpaired bases on the DNA chip were also studied. The pixel intensity of fluorescence of the fragments containing one, two, three, and four unpaired bases were 46,523, 50,364, 33,098 and 30,490, respectively, showing stronger binding of the fusion protein THGLM to the DNA with one or two unpaired bases. Again, the nega-

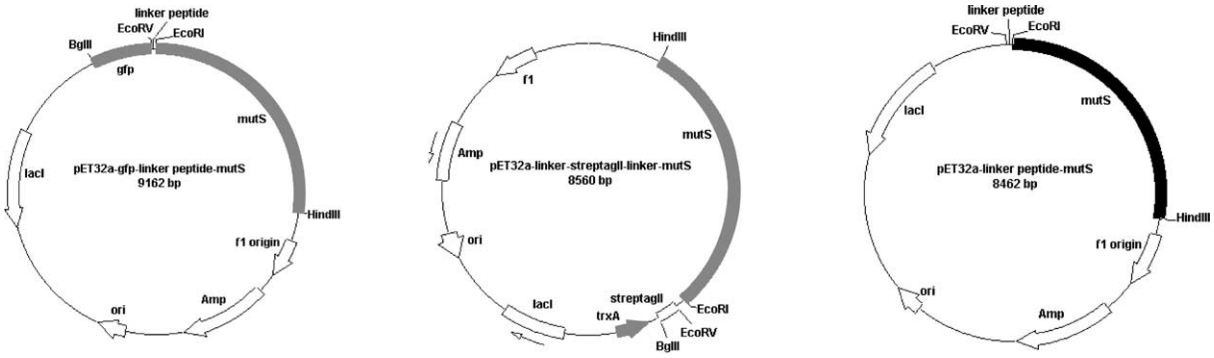

Fig. 1. Structures of expression vectors of the fusion proteins THGLM, THLSLM and THLM.

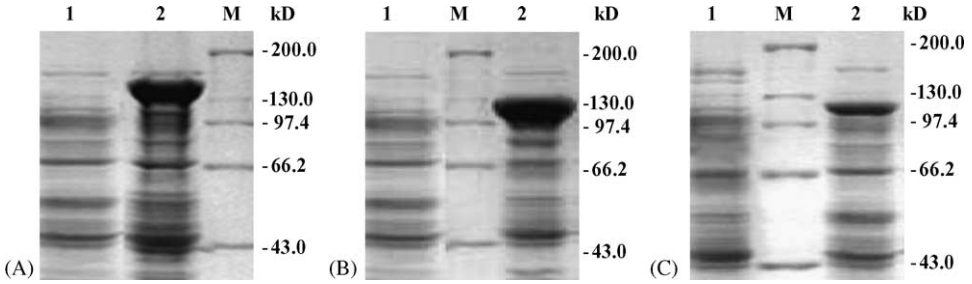

Fig. 2. SDS-PAGE analysis of the fusion proteins. (A) SDS-PAGE analysis of THGLM: 1, the crude protein of AD494(DE3)/pET32a; 2, the crude protein of AD494(DE3)/pET32a-*gfp-Linker-mutS*; M, HMW Marker. (B) SDS-PAGE analysis of THLSLM: 1, the crude protein of AD494(DE3)/pET32a; M, HMW Marker; 2, the crude protein of AD494(DE3)/pET32a-*Linker peptide-Strep tagII-Linker-mutS*. (C) SDS-PAGE analysis of THLM: 1, the crude protein of AD494(DE3)/pET32a; 2, the crude protein of AD494(DE3)/pET32a-*Linker-mutS*; M, HMW Marker.

tive control (with protein only or full complementary DNA) showed very fain signals (Fig. 4B).

THGLM-based DNA chip method was also applied to detect the mutation in *rpoB* gene from rifampin-resistant *M. Tuberculosis*. Tuberculosis is one of the most deadly and common infectious diseases, whose global spread is further complicated by the ubiquitous appearance of drug-resistant strains (Gutierrez et al., 1999). It has been found that resistance to rifampin is highly related to the mutations in *rpoB* gene (Miller et al., 1994), and the codons 531 and 526 are most associated mutation sites. Hence, the position 531 was selected as a mutation model in this experiment. 130-, 320- and 612-bp-long *rpoB* gene fragments were PCR amplified using both wild-type and mutant with A-C substitution at the position 531 as template, respectively. Each PCR product was mixed with the parallel wild-type DNA and then denatured and annealed to form heteroduplex DNA or homoduplex DNA. Then the diluted heteroduplex DNA and homoduplex DNA were, respectively, immobilized on the chip. The fusion protein THGLM was added to the wells and incubated to specially bind mismatched duplex DNA. The pixel intensities of heteroduplexes were 41,224, 36,514 and 35,918, respectively, and those of their parallel homoduplexes were 3750, 5429 and 14,168, respectively (Fig. 4C).

3.3. Detection of DNA mutation with THLSLM based DNA chip

The protocol is similar to that of THGLM experiment, except that the signal was generated through enzymatic transformation of substrate into color precipitate. The binding of THLSLM to heteroduplex DNA and homoduplex DNA immobilized on the chip matrix was performed and then blocked at room temperature for 20 min. The slide was washed with PBS buffer followed by adding the streptavidin-alkaline phosphatase conjugate and its substrate NBT/BCIP

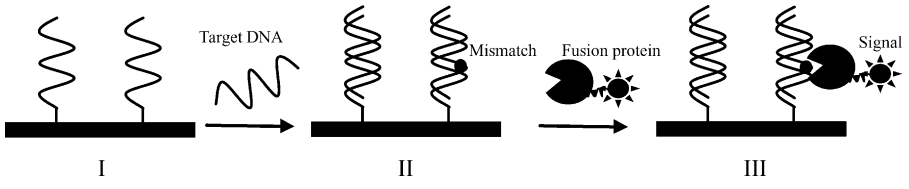

Fig. 3. Schematic representation of MutS-based DNA chip for mutation detection. (I) Probe strands were first immobilized on the glass support; (II) corresponding single-stranded DNA are hybridized to the probe strands, thereby generated heteroduplex DNA and homoduplex DNA; (III) the fusion protein (THGLM, THLSLM and Cy3-labeled THLM) bound preferentially to heteroduplex DNA.

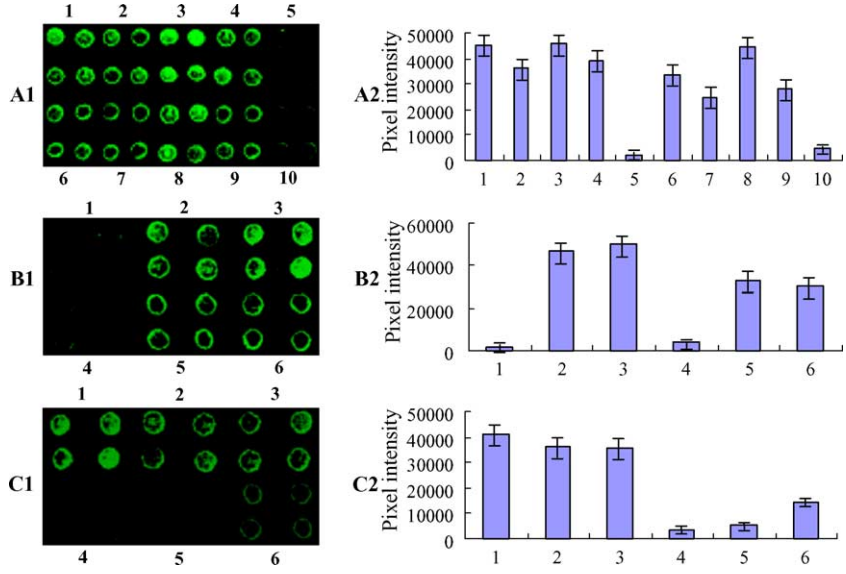

Fig. 4. Mutation detection using THGLM-labeld DNA chip. (A1) Detection images of 30-bp oligonucleotides with different mismatches: 1, a GG mismatch; 2, a TC mismatch; 3, a GT mismatch; 4, a AA mismatch; 5, only protein; 6, a TT mismatch; 7, a CC mismatch; 8, a GA mismatch; 9, a AC mismatch; 10, complementary DNA. (B1) Detection images of 30-bp oligonucleotides with 1–4 unpaired bases: 1, only protein; 2, one unpaired base; 3, two unpaired bases; 4, complementary DNA; 5, three unpaired bases; 6, four unpaired bases. (C1) Detection images of the mutation in *rpoB* gene: 1, 130-bp heteroduplex DNA; 2, 320-bp heteroduplex DNA; 3, 612-bp heteroduplex DNA; 4, 130-bp homoduplex DNA; 5, 320-bp homoduplex DNA; 6, 612-bp homoduplex DNA. (A2), (B2), and (C2) are analysis of the pixel intensity. Sample numbers are corresponding to those of (A1), (B1) and (C1), respectively.

sequentially. The enzymatic reaction was allowed to perform at room temperature for 10 min and then immediately taken photos. Discrimination between mismatched DNA and complementary DNA was evaluated according to the depth of the color development on the spots. The results showed that THLSLM-based DNA chip can readily detect nearly all the mismatches tested, as well as the 30-bp oligonucleotides with 1–4 unpaired bases (Fig. 5A and B). The pixel intensity from the negative controls that contained the fragment of complementary duplex DNA or THLSLM fusion protein only is

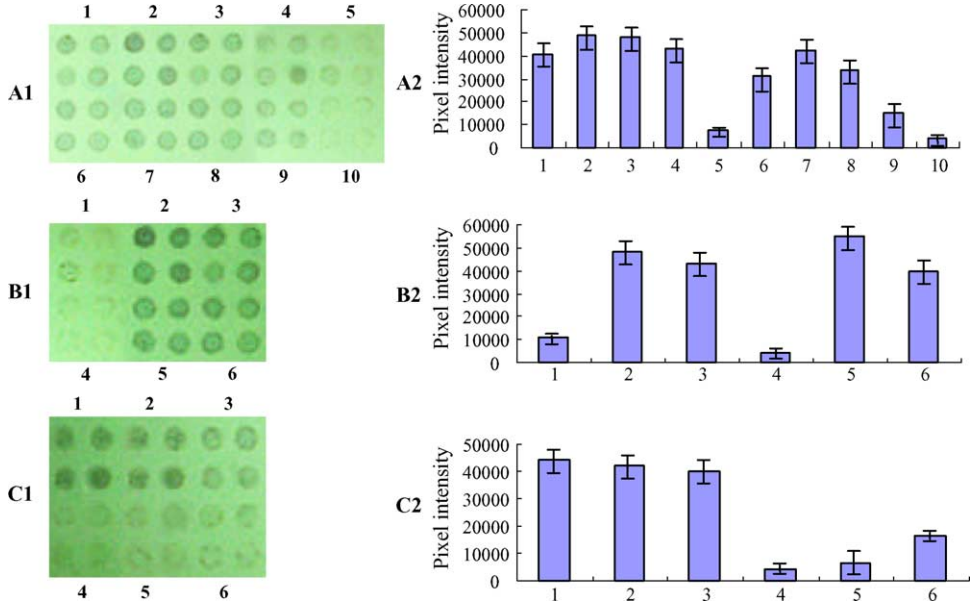

Fig. 5. Detection of DNA mutations using THLSLM-labeled DNA chip. (A1) Detection images of 30-bp oligonucleotides with different mismatches: 1, a TC mismatch; 2, a GT mismatch; 3, a GG mismatch; 4, a AA mismatch; 5, complementary DNA; 6, a AC mismatch; 7, a GA mismatch; 8, a TT mismatch; 9, a CC mismatch; 10, THLSLM only. (B1) Detection images of 30-bp oligonucleotides with 1–4 unpaired bases: 1, complementary DNA; 2, one unpaired base; 3, three unpaired bases; 4, THLSLM only; 5, two unpaired bases; 6, four unpaired bases. (C1) Detection images of the mutation in *rpoB* gene: 1, 130-bp heteroduplex DNA; 2, 320-bp heteroduplex DNA; 3, 612-bp heteroduplex DNA; 4, 130-bp homoduplex DNA; 5, 320-bp homoduplex DNA; 6, 612-bp homoduplex DNA. (A2), (B2), and (C2) are the quantitative analysis of the pixel intensity. Sample numbers are corresponding to those of (A1), (B1) and (C1), respectively.

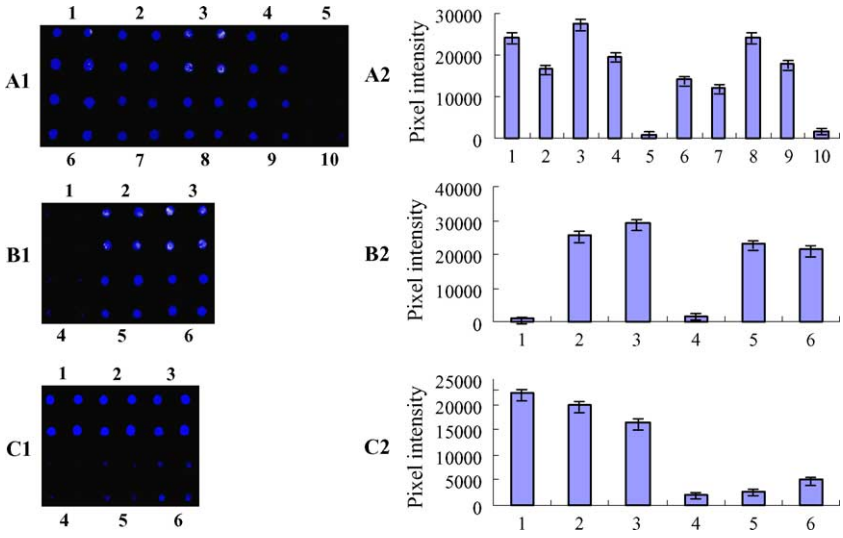

Fig. 6. Detection of DNA mutations using THLM-Cy3 labeled DNA chip. (A1) Detection images of 30-bp oligonucleotides with different mismatches: 1, a GG mismatch; 2, a TC mismatch; 3, a GT mismatch; 4, a AA mismatch; 5, THLM-Cy3 only; 6, a TT mismatch; 7, a CC mismatch; 8, a GA mismatch; 9, a AC mismatch; 10, complementary DNA. (B1) Detection images of 30-bp oligonucleotides with 1–4 unpaired bases: 1, THLM-Cy3 only; 2, one unpaired base; 3, two unpaired bases; 4, complementary DNA; 5, three unpaired bases; 6, four unpaired bases. (C1) Detection images of the mutation in *rpoB* gene: 1, 130-bp heteroduplex DNA; 2, 320-bp heteroduplex DNA; 3, 612-bp heteroduplex DNA; 4, 130-bp homoduplex DNA; 5, 320-bp homoduplex DNA; 6, 612-bp homoduplex DNA. (A2), (B2), and (C2) are quantitative analysis of the pixel intensity. Sample numbers are corresponding to those of (A1), (B1) and (C1), respectively.

7143 and 4375, respectively. It showed that they also produced faint color change, which, however, was obviously weaker than that of mismatched or unpaired DNA, the pixel intensity of which is between 15,409 and 54,549. Results from the variance analysis and time-test also showed that the color among different mismatches had no notable discrimination, but the difference between mismatched duplex DNA and complementary duplex DNA was remarkable.

The THLSLM-based DNA chip method was also applied to detect the mutation in *rpoB* gene (Fig. 5C). The result was very similar to that of using THGLM system. The pixel intensities of heteroduplexes were 44,165, 42,418 and 39,720, respectively, for 130-, 320- and 612-bp- long *rpoB* fragments, while those of their parallel homoduplexes were 4387, 6290 and 16,072.

3.4. Mutation detection of oligonucleotides with THLM-Cy3 labeled DNA chip

Fig. 6 shows the experiment results, which can discriminate between mismatched and complementary DNA. The pixel intensity of fluorescence of the spots containing different mismatches was between 27,691 and 12,032 with the background signals varying from 1732 to 912 (Fig. 6A). The pixel intensity of fluorescence of the fragments containing one, two, three, and four unpaired bases were 25,760, 29,412, 22,891 and 21,337, respectively (Fig. 6B), showing stronger binding of the fusion protein Cy3-labeled THLM to the DNA with one or two unpaired bases. Again, the negative control showed very faint signals and their pixel intensity were 1062 and 1659. The results from detection of the mutation in the *rpoB* gene showed that the pixel intensities of heteroduplexes were 22,147, 19,820 and 16,467, respectively, and those of their parallel homoduplexes were 1918, 2577 and 5183 in the same sequence of 130, 320 and 612 bp, respectively (Fig. 6C).

4. Discussion

In the recombination of MutS fusions, thioredoxin and his.tag were used as fusion partners in order to circumvent the formation of inclusion bodies caused by over expression of heterologous gene and facilitate the purification of the desired protein, respectively (Dickason et al., 1995). A flexible linker peptide composed of serine and glycine was designed to minimize the steric hindrance and provide enough space for the components of the fusion proteins to remain their native conformation (Shao et al., 2000; Zhou et al., 2001). With this linker, the biological activities of MutS and its fusion partners (thioredoxin, his.tag, GFP, and Strep tagII, etc.) remained perfectly.

The optimal temperature for *E. coli* growth is 37 °C. Expression of the recombinant protein at this temperature is often preferred to keep growth periods as short as possible. However, empirical results have shown that a number of recombinant proteins are primarily produced in an insoluble, i.e. nonfunctional, form at 37 °C. The yield of soluble, functional protein can be substantially increased in most cases by lowering the initial growth temperature to between 22 °C and 30 °C. Consequently, in the present study, combining the growth of the cells with the expression of the recombinant protein, 28 °C was chosen to perform the expression of the

Table 2
Characters of the DNA chips using different MutS fusion labeling system

Labeling system	Signal to noise[a]	Detection limit[b] (μM)	Operating step[c]	Operating time (h)	Cost
THGLM	10/1	0.04–0.5	6	4.5	++
THLSLM	6/1	0.12–0.1	7	5.5	+
THLM-Cy3	15/1	0.02–1	6	4	+++

[a] Calculated by the discrimination signal ratio between the mean value of the impaired or unpaired DNA and that of the control.
[b] It is the concentration range of detected DNA sample, in which there are producing the discrimination signal on the chip, but not including PCR process.
[c] It includes the steps which begin from the process on the chip to the detection of the signal, and not includes PCR process and other pretreatment.

fusion protein THGLM, and 30 °C was chosen for the expression of the fusion proteins THLM and THLSLM.

The amount of the fusion proteins in the mismatch-binding reaction was found to significantly affect the results of the assay. It was found that MutS fusion protein with high concentration bound small amount of complementary duplex DNA in the previous study (Bi et al., 2003).

Mutation detection of oligonucleotides showed that the three fusion proteins bound to nearly all the duplex DNA with mismatched base pair and 1–4 unpaired bases. It is obviously that the GT mismatch generates the brightest signal and the weakest mismatch binding was observed to be a CC mismatch, the latter mismatch has been found to be generally refractory to repair by the *E. coli* MMR system (Modrich, 1991). The fragment of Complementary duplex DNA also produced faint signal, which, however, was much weaker than that of a CC mismatch. Mismatched duplexes are easily distinguished from complementary duplexes. However, the mismatches cannot be readily distinguished from one another because most of the mismatches generated approximately the same signal level. The weak signal level for a CC mismatch does not diminish the utility of these approaches for mutation detection, since every wild-type and mutant pairing gives rise to two different mismatches (GG and CC). The MutS-based DNA chip could bind GG mismatch well and then give strong signal. The mechanism of the discrepancy of binding affinity is not very clear but believed due to the structural differences between these base pairs (Lamers et al., 2000). Taking GT mismatch as an example, the hypothesis is that three major interactions occur between MutS and mismatched duplex DNA that allow the recognition of the mismatched bases. The amino acid Glu38 of the mismatch-binding domain of MutS forms hydrogen bonds to G10, and to T22 of the mismatch. Asp35 of MutS forms a hydrogen bond with G9 of the mismatch. The third interaction is the wedging of Phe36 of MutS into the duplex DNA, which stacks with the T22 mismatch. The Phe36 will only wedge in where the DNA backbone is flexible due to a mismatch or looped out base (Malkov et al., 1997). Other mismatches (GG, AA, GA, CC) should have different interaction mechanisms. It must be interesting to take detailed investigations on them at super molecular structural level.

One interesting phenomenon in detection of mutation of *rpoB* gene is that the signal level from high to low was in sequence of 130-, 320- and 612-bp fragments, but the noise (background signal) level of the homoduplexes of the same fragments increased with the increase of DNA length (Figs. 4C, 5C and 6C). Possible explanation is that the MutS fusion is more favor to the shorter mismatched DNA fragments, and the longer DNA probes might form mismatches when annealed with the genomic DNA from the homologous chromosome or generate some secondary structure with mismatches, causing non-specific bindings. This finding suggests that the amount of the MutS fusion proteins in the binding reaction mixture must be carefully controlled. Under the controlled experiment condition, these non-specific signals were much weaker than those of the specific signals and thus did not affect the results.

In the present study, three different MutS fusion protein were constructed and used as the recognition elements of the DNA chip for discrimination of DNA mutation, and the experiment data showed they all functioned well. Table 2 is the summarization of the performance of these MutS fusion protein labeled DNA chips. Generally, THGLM protocol is attractive not only for the high stability of the green fluorescent chromophore of GFP, but also for the increased sensitivity because each MutS contains two GFP molecules, as MutS is a dimmer and GFP is a monomer. In addition, the system requires less expensive instrument such as fluorescence microscopy and can be applied to detect the DNA mutations conveniently and rapidly with less background noise.

The Cy3-labeled THLM system showed the highest sensitivity among all three fusions. There was almost no background signal, the ratio of signal to noise is 15/1. Beside, the method is simple and rapid, allowing the batch operation for large number of samples, although the expensive fluorescent scanning instrument is required.

Compared to the other two MutS fusions, the main deficiency of the THLSLM fusion protein is its complex operation protocol, which needs the post incubations for Strep tagII-streptavidin interaction and enzymatic reaction after binding of THLSLM to the mismatched DNA on the DNA chip. These two additional steps prolong the detection time and require skills. Another drawback of the system is that the signal to noise ration is lower than that of other two systems, although the generated response signal to the DNA mutation is always distinct. However, the method has its unique advantage: the detection result can be viewed by the naked eyes without aid of any instruments. This feature provides the more opportunities to allow the application of the method, either preliminary screening or routine detection of DNA mutations.

5. Conclusion

We proposed three fusion systems for detection of gene mutations based on the principle of MutS recognizing and binding mismatched DNA, and demonstrated them experimentally. The success of three methods largely relies on the application of the gene manipulation technology, which ensured that the components of the fusion proteins THGLM, THLSLM and THLM retain their native bioactivities. The preliminary study has been demonstrated that the MutS-based DNA chips feature high specificity, wide range of target DNA size (30–612 bp), ease of handling, and short detection cycle. The established approaches have great potential in multiplex/parallel detection of gene mutations, and could be used in high throughput preliminary screen of gene mutations.

Acknowledgements

We sincerely thank the supports from the National Science Foundation of China (No. 30270306) Ministry of Science & Technology and the Chinese Academy of Science (No. kscx1-06-01).

References

Beaulieu, M., Larson, G.P., Geller, L., Flanagan, S.D., Krontiris, T.G., 2001. PCR candidate region mismatch scanning: adaptation to quantitative, high-throughput genotyping. Nucleic Acids Res. 1, 1114–1124.

Behrensdorf, H.A., Pignot, M., Windhab, N., Kappel, A., 2002. Rapid parallel mutation scanning of gene fragments using a microelectronic protein-DNA chip format. Nucleic Acids Res. 15, e64.

Beltran-Valero De Bernabe, D., Currier, S., et al., 2002. Mutations in the O-mannosyltransferase gene POMT1 give rise to the severe neuronal migration disorder Walker-Warburg syndrome. J. Hum. Genet. 71 (5), 1033–1043.

Bi, L.J., Zhou, Y.F., Zhang, X.E., Deng, J.Y., Zhang, Z.P., Xie, B., Zhang, C.G., 2003. A MutS-based protein chip for detection of DNA mutations. Anal. Chem. 75, 4113–4119.

Deeble, V.J., Roberts, E., Robinson, M.D., Woods, C.G., Bishop, D.T., Taylor, G.R., 1999. Comparison of enzyme mismatch cleavage and chemical cleavage of mismatch on a defined set of heteroduplexes. Genet. Testing 1, 253–259.

Dickason, R.R., Edwards, R.A., Bryan, J., Huston, D.P., 1995. Versatile *E. coli* thioredoxin specific monoclonal antibodies afford convenient analysis and purification of prokaryote expressed soluble fusion protein. J. Immunol. Methods 185, 237–244.

Ellis, L.A., Taylor, G.R., Banks, R., Baumberg, S., 1994. MutS binding protects heteroduplex DNA from exonuclease digestion in vitro: a simple method for detecting mutations. Nucleic Acids Res. 11, 2710–2711.

Geschwind, D.H., Rhee, R., Nelson, S.F., 1996. A biotinylated MutS fusion protein and its use in a rapid mutation screening technique. Genet. Anal. 13, 105–111.

Gotoh, M., Hasebe, M., Ohira, T., Hasegawa, Y., Shinohara, Y., Sota, H., Nakao, J., Tosu, M., 1997a. Rapid method for detection of point mutations using mismatch binding protein (MutS) and an optical biosensor. Genet. Anal. 14, 47–50.

Gotoh, M., Hasebe, M., Ohira, T., Tosu, M., 1997b. Gene diagnosis with an affinity sensor, BIACORE–principle and applications. Rinsho Byori 45, 224–228.

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. 16, 535–540.

Guldberg, P., Guttler, F., 1993. A simple method for identification of point mutations using denaturing gradient gel electrophoresis. Nucleic Acids Res. 21, 2261–2262.

Gutierrez, M.C., Galan, J.C., Blazquez, J., Bouvet, E., Vincent, V., 1999. Molecular markers demonstrate that the first described multidrug-resistant Mycobacterium bovis outbreak was due to *Mycobacterium tuberculosis*. J. Clin. Microbiol. 37 (4), 971–975.

Hacia, J.G., 1999. Resequencing and mutational analysis using oligonucleotide microarrays. Nat. Genet. 21, 42–47.

Han, A., Shibata, T., Takarada, T., Maeda, M., 2002. Gene mutation assay using a MutS protein-modified electrode. Nucleic Acids Res. Suppl. 2, 287–288.

Howard, H.C., Mount, D.B., et al., 2002. The K-Cl cotransporter KCC3 is mutant in a severe peripheral neuropathy associated with agenesis of the corpus callosum. Nat. Genet. 32 (3), 384–392.

Jiricny, J., Su, S.S., Wood, S.G., Modrich, P., 1988. Mismatch-containing oligonucleotide duplexes bound by the *E. coli* mutS-encoded protein. Nucleic Acids Res. 25, 7843–7853.

Lamers, M.H., Perrakis, A., Enzlin, J.H., Winterwerp, H.H., de Wind, N., Sixma, T.K., 2000. The crystal structure of DNA mismatch repair protein MutS binding to a G × T mismatch. Nature 407, 711–717.

Lieb, M., 1987. Bacterial genes mutL, mutS, and dcm participate in repair of mismatches at 5-methylcytosine sites. J. Bacteriol. 169, 5241–5246.

Lishanski, A., Ostrander, E.A., Rine, J., 1994. Mutation detection by mismatch binding protein, MutS, in amplified DNA: application to the cystic fibrosis gene. Proc. Natl. Acad. Sci. U.S.A. 91, 2674–2678.

Lu, A.L., Clark, S., Modrich, P., 1983. Methyl-directed repair of DNA base-pair mismatches in vitro. Proc. Natl. Acad. Sci. U.S.A. 80, 4639–4643.

Malkov, V.A., Biswas, I., Camerini-Otero, R.D., Hsieh, P., 1997. Photocross-linking of the NH_2-terminal region of Taq MutS protein to the major groove of a heteroduplex DNA. J. Biol. Chem. 272, 23811–23817.

McCready, M.E., Sweeney, E., et al., 2002. A novel mutation in the IHH gene causes brachydactyly type A1: a 95-year-old mystery resolved. Hum. Genet. 111, 368–375.

Miller, L.P., Crawford, J.T., Shinnick, T.M., 1994. The *rpoB* gene of *Mycobacterium tuberculosis*. Antimicrob. Agents Chemother. 38 (4), 805–811.

Modrich, P., 1991. Mechanisms and biological effects of mismatch repair. Annu. Rev. Genet. 25, 229–253.

Parsons, B.L., Heflich, R.H., 1997. Evaluation of MutS as a tool for direct measurement of point mutations in genomic DNA. Mutat. Res. 21, 277–285.

Parsons, B.L., Heflich, R.H., 1998. Detection of basepair substitution mutation at a frequency of 1×10 (-7) by combining two genotypic selection methods, MutEx enrichment and allele-specific competitive blocker PCR. Environ. Mol. Mutagen. 32, 200–211.

Sachadyn, P., Stanisawska, A., Kur, J., 2000. One tube mutation detection using sensitive fluorescent dyeing of MutS protected DNA. Nucleic Acids Res. 15, E36.

Saiki, R.H., Sharf, S., et al., 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230, 1350–1354.

Schaller, A., Guo, M., Gisanrin, O., Zhang, Y., 2002. *Escherichia coli* genes involved in resistance to pyrazinoic acid, the active component of the tuberculosis drug pyrazinamide. FEMS Microbiol. Lett. 4, 265–270.

Shao, W.H., Zhang, X.E., Liu, H., Zhang, Z.P., Cass, A.E., 2000. Anchor-chain molecular system for orientation control in enzyme immobilization. Bioconjug. Chem. 11, 822–826.

Smith, J., Modrich, P., 1996. Mutation detection with MutH, MutL, and MutS mismatch repair proteins. Proc. Natl. Acad. Sci. U.S.A. 30, 4374–4379.

Sun, H.B., Yokota, H., 2000. MutS-mediated detection of DNA mismatches using atomic force microscopy. Anal. Chem. 15, 3138–3141.

Sugano, K., Fukayama, N., Ohkura, H., Shimosato, Y., Yamada, Y., Inoue, T., Sekiya, T., Hayashi, K., 1995. Single-strand conformation polymorphism analysis by perpendicular temperature-gradient gel electrophoresis. Electrophoresis 16, 8–10.

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. 15, 489–494.

Tanigawa, M., Gotoh, M., Machida, M., Okada, T., Oishi, M., 2000. Detection and mapping of mismatched base pairs in DNA molecules by atomic force microscopy. Nucleic Acids Res. 1, E38.

Wagner, R., Dean, A., 2000. The use of immobilized mismatch binding protein in mutation/SNP detection. Methods Mol. Biol. 152, 159–168.

Wagner, R., Debbie, P., Radman, M., 1995. Mutation detection using immobilized mismatch binding protein (MutS). Nucleic Acids Res. 11, 3944–3948.

Wong, C., Dowling, C.E., et al., 1987. Characterization of beta-thalassaemia mutations using direct genomic sequencing of amplified single copy DNA. Nature 330, 384–386.

Zhou, Y.F., Zhang, X.E., Liu, H., Zhang, Z.P., Zhang, C.G., Cass, A.E., 2001. Construction of a fusion enzyme system by gene splicing as a new molecular recognition element for a sequence biosensor. Bioconjug. Chem. 12, 924–931.