Analyst

Pathogen-specific DNA sensing with engineered zinc finger proteins immobilized on polymer chip

Journal:	*Analyst*
Manuscript ID	AN-COM-03-2018-000395.R2
Article Type:	Communication
Date Submitted by the Author:	29-Jun-2018
Complete List of Authors:	Ha , Dat ; Western Kentucky University, Chemistry ghosh, Sthitodhi; University of Cincinnati, College of Engineering and Applied Science; Ahn, Chong ; University of Cincinnati, College of Engineering and Applied Science Segal, David; University of California, Genome Center and Department of Pharmacology Kim, Moon-Soo; Western Kentucky University, Chemistry

Pathogen-specific DNA sensing with engineered zinc finger proteins immobilized on polymer chip

Dat Thinh Ha[1], Sthitodhi Ghosh[2], Chong H. Ahn[2], David, J. Segal[3], Moon-Soo Kim[1,*]

[1]Department of Chemistry, Western Kentucky University, Bowling Green, Kentucky 42101, USA

[2]Microsystem and BioMEMS Laboratory, University of Cincinnati, Cincinnati, Ohio 45221, USA

[3]Genome Center, Department of Biochemistry and Molecular Medicine, University of California, Davis, CA 95616, USA.

*To whom correspondence should be addressed. Email: moon-soo.kim@wku.edu; Tel: +1-270-745-4362; Fax: +1-270-745-5361

Abstract

A specific double-stranded DNA sensing system is of great interest for diagnostic and other biomedical applications. Zinc finger domains, which recognize double-stranded DNA, can be engineered to form custom DNA-binding proteins for recognition of specific DNA sequences. As a proof of concept, a sequence-enabled reassembly of TEM-1 β-lactamase system (SEER-LAC) was previously demonstrated to develop zinc finger protein (ZFP) arrays for the detection of a double-stranded bacterial DNA sequence. Here, we implemented the SEER-LAC system to demonstrate the direct detection of pathogen-specific DNA sequences present in *E. coli* O157:H7 on the lab-on-a chip. ZFPs custom-designed to detect shiga toxin in *E. coli* O157:H7 were immobilized on the cyclic olefin copolymer (COC) chip, which can function as a non-PCR based molecular diagnostic. Pathogen-specific double-stranded DNA was directly detected by engineered ZFPs immobilized on the COC chip with high specificity, providing a detection limit of 10 fmole of target DNA in colorimetric assay. Therefore, in this study, we demonstrated a great potential of ZFP arrays on the COC chip for further development of a simple and novel lab-on-a chip technology for detection of pathogens.

Introduction

DNA detection technologies play an important role in diagnostic applications in the areas of publich health and biomedicine.[1] Although numeuros methods for DNA detection have been developed, a simple, sensitive and rapid technology for the detection of pathogen-specific double-stranded (ds) DNA sequences still remains a challenge in pathogen detection and clinical diagnostics. DNA diagnostics require a detection method with a signal transducer.[2] Most of the current methods are based on DNA denaturation and subsequent hybridization with its complimentary probe, such as polymerase chain reaction (PCR), DNA microarray, and fluorescence *in situ* hybridization (FISH). PCR has provided a sensitive and faster method for pathogen detection than a traditional culture-based method.[3-5] On the other hand, nucleic acid amplification by PCR also requires multiple primers and precise thermal cycling conditions to be discriminated from non-specific amplification. DNA microarray allows one to identify multiple pathogens simultaneously, but it requires DNA labeling and hybridization with their complimentary probes, which can increase the inaccuracy due to the cross-reacting of several probes with incorrect targets.[6] FISH has been studied in biomedical/clinical researches, especially in diagnostics for visualizing and detecting specific DNA/RNA sequences.[7] However, FISH can be time-consuming with limited sensitivity and its standardized protocol is not widely available. Thus, direct detection of specific dsDNA through DNA-binding proteins enables us to avoid the need for additional laborious steps involved in DNA denaturation and subsequent hybridization. This is the novel aspect of our study which would enable us to develop a simple and rapid technology for DNA detection.

A Cys2-His2 zinc finger (ZF) domain, which contains 30 amino acids, is one of the most common DNA-binding domains that can be found in various eukaryotic genomes.[8, 9] A ZF

domain folds into a ββα structure that is stabilized by zinc coordination and hydrophobic residues.[10, 11] Each domain can recognize three to four DNA nucleotides. Multiple zinc finger domains can be linked together to form multi-finger proteins to recognize extended DNA sequence.[10, 12] Construction of multi-finger proteins enables us to improve the binding affinity and specificity.[10, 12] To further modify the specificity of zinc finger proteins (ZFPs), modular assembly approach has been used to assemble ZF domains targeting their respective 3 bp subsites to recognize specific sequences of an interest. [12-15] Theoretically, a six ZFP can recognize 18 bp of a specific DNA sequence, sufficient enough to recognize a unique site within all known genomes.[12, 13, 16] Therefore, customized ZFPs can be created to detect virtually any DNA sequence.[9, 17]

Previously, a colorimetric detection method for visualizing specific DNA detection, called the SEquence-Enabled Reassembly of β-lactamase (SEER-LAC) system, has been developed.[9, 17] This method utilizes enzymatic activities of TEM-1 β-lactamase which will hydrolyze the nitrocefin substrate, creating a visible color change from yellow to red.[9] In this system, β-lactamase was dissected into two inactive fragments (LacA and LacB), each linked with a ZFP.[9] Upon specific binding of ZFPs to their target DNA sequence, bringing the two inactive fragments in a close proximity, LacA and LacB would reassemble to become a full-length β-lactamase, restoring its enzymatic activities.[9] Further applications of the SEER-LAC system has led to the development of ZFP array on a poly(ethylene glycerol) (PEG) hydrogel-coated glass slide to detect bacterial dsDNA sequences.[17]

Here, we implemented the SEER-LAC system to demonstrate the direct detection of pathogen-specific DNA sequences present in *E. coli* O157:H7 on a ZFP array on a transparent polymer surface. Engineered ZFPs were deposited on a cyclic olefin copolymer (COC) chip,

which can function as a simple and inexpensive detection platform for specific pathogen detection. Our approach provides rapid visual detection along with high specificity and sensitivity, suggesting that ZFP arrays on the COC chip could be further developed into a novel and reliable molecular diagnostic device for multiplexed detection of pathogens.

Experimental

Construction, expression and purification of stx2 ZFPs

ZFPs were constructed by the modular assembly method using the Barbas set of pre-defined ZF modules.[15] The DNA coding regions for each ZFP were commercially synthesized by Bio Basic. Each pair of ZFPs consists of ZFP(A) and ZFP(B). The ZFP(A)s (stx2_268, stx2_560, and stx2_1093) were sub-cloned between the XmaI and HindIII sites of pMAL-c2X LacA-rrsA1175, replacing the C-terminal rrsA1175 ZFP. The ZFP(B)s (stx2_233, stx2_525, and stx2_1128) were sub-cloned between the BamHI and AgeI sites of pMAL-c2X rrsA1192-LacB, replacing the N-terminal rrsA1192 ZFP. The pMAL vector was used for bacterial expression of the proteins as fusions with an N-terminal maltose binding protein (MBP) as a purification tag. Proteins were expressed in *E. coli* BL21 (Invitrogen) upon induction with 1 mM isopropyl b-D-1-thiogalactopyranoside (IPTG) at an OD_{600} of 0.6–0.8 for 3 h at 37^{o}C. Cells were pelleted and re-suspended in Zinc Buffer A (ZBA: 100 mM Tris base, 90 mM KCl, 1 mM $MgCl_2$ and 100 mM $ZnCl_2$ at pH 7.5) including 5 mM dithiothreitol (DTT) and 50 mg/ml RNase A. After sonication, proteins in cell lysates were applied to an amylose resin column (Bio-rad) pre-equilibrated with ZBA + 5 mM DTT, washed with ZBA + 2 M NaCl and ZBA + 1 mM Tris(2-carboxyethyl)phosphine (TCEP), and eluted in ZBA + 10 mM Maltose + 1 mM TCEP.

Concentration and purity were assessed by Coomassi-stained polyacrylamide gel electrophoresis with sodium dodecyl sulfate (SDS–PAGE) and Bradford assay using bovine serum albumen (BSA) standards. Purified protein was stored on ice at 4^{o}C until use.

Fabrication of the COC chip

Cyclic Olefin Copolymer (COC) of grade 5013-S was obtained from Topas Advanced polymers (Kentucky, USA). COC chips were fabricated using a polymer injection molding machine (BOY 22A, Procan, CT). The COC thermoplastic in the form of pellets is heated in a barrel and this molten plastic in the barrel is forced into a small cavity containing the features to be replicated. The cavity is cooled down and the solid plastic part is ejected from the mold cavity. As per the processing datasheet of Topas 5013S-04, the injection molding parameters were optimized and are as follows: nozzle temperature 565 °F, mold temperature 480 °F, speed of injection 85 mm/sec, and back pressure 2000 psi. The process cycle time was optimized to 35 seconds thus ensuring lower injection molding cost and increasing the throughput of the process. The COC chip was fabricated as an array spot lab-chip. Plain COC disks of 1 mm thickness and 76.2 mm diameter were produced by injection molding using a blank polished Aluminum disk as the mold. An array spot aligned with an individual well in a 96-well plate was marked and was covered with a chemical resistance tape. The spots were 7 mm in diameter with a pitch of 9 mm, same as in 96 well plates (Fig. 1(A)). The rectangular array lab chips were incubated in Phosphate buffered Saline (PBS) for 15-45 mins and then dried. The tapes were removed and the chips were ready for testing.

ZFP array and nitrocefin assay

DNA target oligonucleotides were commercially synthesized by IDT and prepared by heating to 95°C for 10 minutes, then slowly cooling to room temperature to form hairpins containing a four-thymidine loop. The sequences of hairpin DNA target oligonucleotides are provided in the Supplementary Data (Figure S1). A silicone gasket with a diameter of 6 mm and a well depth of 1 mm (Grace Bio-Labs, Bend, OR) was placed onto the cyclic olefin copolymer (COC) surface to confine the areas of ZFP immobilization/enzymatic reaction before arraying the ZFPs. 5 μL of a purified protein LacA-ZFP(A) at a concentration of 2.5 μM was pipetted onto the COC surface and incubated for 40 minutes. 10 μL of hairpin target DNA solution was added on the ZFP array and incubated for 20 minutes to allow DNA binding to the ZFP. The slide was washed with ZBA + 50 mM KCl and ZBA + 0.05% Tween-20, followed by air-drying. 10 μL of a purified protein ZFP(B)-LacB was added on the ZFP array and incubated for 20 minutes to allow the ZFP to bind the DNA that was complexed with the LacA-ZFP(A). The slide was washed with ZBA + 50 mM KCl and ZBA + 0.05% Tween-20, followed by air-drying. After placing the slide onto a 96-well plate and aligning the arrays with the wells, 20 μL of 1 mM nitrocefin (Calbiochem, San Diego, CA, USA) was added to the ZFP array. Absorbance at 486 nm was monitored with a Spectramax 190 (Molecular Devices, Sunnyvale, CA, USA). All experiments were repeated in duplicate, and the standard error was calculated from duplicate samples.

Electrophoretic mobility shift assay (EMSA)

Complementary pairs of 5'-biotin labeled forward and 5'-poly T reverse oligonucleotides were annealed by heating to 95°C for 3 minutes, and cooling to 4°C by 1°C per 50 seconds to obtain double-stranded target DNAs. Binding reactions were performed at room temperature in the dark for 1.5 hour in ZBA containing 150 mM KCl, 5 mM DTT, 10% glycerol, 0.1 mg/ml BSA, 0.05% NP-40, 5 pmol target DNA and purified ZFPs with a concentration of 0.6–500 nM. Gel

electrophoresis was performed in the cold room on a 10% native polyacrylamide gel in 0.5 X TBE buffer. After blotting on a nylon membrane by transferring in the cold room, the DNA was cross-linked by a UV cross-linker for 4 minutes. After that, EMSA was performed using the Light Shift Chemiluminescent EMSA Kit (Pierce, Rockford, IL, USA) according to the manufacturer's protocol. Chemiluminescent signal was read using AlphaImager HP (ProteinSimple, San Jose, CA, USA).

Binding site selection assay using Bind-n-Seq

Binding sites of engineered ZFPs were identified using Bind-n-Seq as described.[18] Barcoded 95-mer double-stranded oligonucleotides containing Illumina primer binding sites and a 21-nt random region were incubated with different protein concentrations (500 nM, 50 nM, and 5 nM final) and salt concentrations (100 mM, 50 mM, and 1 mM final) in Bind-n-Seq (BnS) buffer (0.12 μg/μL herring sperm DNA, 5 mM DTT, 1% BSA). Binding reactions were carried out at room temperature for 2 hours. Bound complexes were precipitated using amylose resin and enriched by six washing steps with the corresponding salt buffers. Eluted DNA were quantified and sufficiently amplified for sequencing on an Illumina sequencer. Sequencing reads were filtered and sorted using custom Perl scripts found in the MERMADE package, an updated version of the Bind-n-Seq data analysis pipeline. MERMADE is freely available with user documentation at http://korflab.ucdavis.edu/Datasets/BindNSeq.

Results and discussion

ZFP array on the COC chip

Previously, we have demonstrated our initial proof-of concept on a simple capture-detection probe assay to detect the 16S ribosomal DNA sequence of a non-pathogenic strain of *E. coli*.[17] In this study, we implemented our system to further develop a molecular diagnostic for detecting a food-borne pathogen *E. coli* O157:H7. In addition, a new copolymer surface was used to immobilize ZFPs as a choice of a substrate material for the lab-on-a chip. Fig. 1 represents our system where a capture probe ZFP(A) linked to a LacA fragment was immobilized on the COC chip, followed by dsDNA binding with the protein, forming the protein-DNA complex. Subsequently, a washing step was performed to wash off unbound molecules, followed by the binding of a detection probe ZFP(B) attached to a LacB fragment. When brought into a close proximity upon ZFPs binding to the target sites, the two inactive fragments LacA and LacB would reassemble into a full-length β-lactamase, which then hydrolyzes the β-lactam ring of nitrocefin substrate, converting its color from yellow to red. This system generated a visual signal that indicates ZFPs binding to their specific target DNA sequences.

Three pairs of ZFPs were engineered to recognize a pathogen-specific gene *stx2* encoding for Shiga toxin present in *E. coli* O157:H7. Three pairs of six-finger ZFPs were constructed targeting three different sites in *stx2* (Supplementary data), which are stx2_233/stx2_268, stx2_525/stx2_560, and stx2_1093/stx2_1128 (Table 1). The Lac A was attached to the N-terminus of the capture probe ZFP(A)s and LacB to the C-terminus of the detection probe ZFP(B)s. Since stx2_1093/stx2_1128 pair binds to its target DNA on the sense strand, LacA was linked to stx2_1093, and LacB was linked to stx2_1128. The other two pairs, however, bind to their target DNA on the antisense strand. Thus, LacA was linked to stx2_268 and stx2_560, while LacB was linked to stx2_233 and stx2_525. This design would ensure LacA and LacB fragments to be justaposed upon a pair of ZFPs binding to the continuous target sites of 36 bp.

Several polymers including COC are considered for a substrate material for the lab-on-a-chip application.[19] An ideal point-of-care diagnostic system should consist of disposable cheap cartridges which can be manufactured in large quantities at a low cost. The thermoplastic polymer COC has a definite advantage over PEG-coated glass slides in this regard. Using a single Aluminum mold, hundreds of replicable COC chips can be made, which makes this polymer ideal for lab-on-a-chip-based diagnostic applications. Moreover, different types of microchannel surfaces can be more easily fabricated on COC using the injection molding technique. This can provide larger surface-to-volume ratio and improved sensitivity in future application. Also, COC exhibits optical transparency over a wide range including the UV spectrum which makes it ideal for biochemical analysis and bio-optical application.[20, 21] In addition to that, resistance to polar solvents, high biological compatibility, low background noise, and easy control of non-specific adsorption prompt the use of COC as a better substrate material in this study.[20-22] COC displays a very high flow rate during injection molding as compared to other polymer materials like PC (polycarbonate) or PMMA (polymethyl methacrylate) and its low viscosity allows for lower injection pressure and better fills.[23] Thus, COC can serve as an ideal platform for developing low-cost, disposable lab-on-chip devices for pathogen detection. In this study, the COC chip was fabricated to function as a surface for ZFP immobilization. The capture probe LacA-ZFP(A) was deposited within a confined area of a silicone gasket on the COC chip.

Sensitivity

ZFP arrays were performed with each pair of ZFPs at various target DNA concentrations ranging from 2.5 μM to 1 nM to determine the sensitivity of our system in the presence of their own target DNAs (Fig. 2). A linear DNA dose calibration curve was generated from the data,

showing a linear and quantitative assay. All the ZFP pairs were able to detect their own target DNA in the range of 5 nM to 2.5 μM. Among the three pairs, the ZFP pair LacA stx2_268 and stx2_233 LacB was most sensitive because it was able to detect 1 nM of DNA which is equal to 280 pg or 10 fmol of oligonucleotide target DNA ($P< 0.05$). The other two pairs of ZFPs were sensitive enough to detect 5 nM of target DNA, but not 1 nM of target DNA, indicating a limit of detection > 1 nM ($P< 0.05$).

The limit of detection of this study is ≤10 fmole, which is equivalent to 280 pg of DNA. Using a new surface platform resulted in a five-fold improvement in the sensitivity compared to the previous study.[17] For *E. coli* O157, real-time PCR provides the limit of detection of 2×10^2 CFU/ml with an assay time of 24 hours.[24] In oligonucleotide DNA microarray, the limit of detection for *E. coli* O157 is 0.1 pg for each genomic DNA.[25] Although the detection methods of the leading DNA techniques are different from our system, our method is currently not as sensitive as these methods because we are still at the early stage of further developing our sensing system to improve the sensitivity. More importantly, the novel aspect of our system is that it does not require DNA labeling or DNA denaturation and subsequent hybridization under controlled conditions which the leading DNA-based methods require. In addition, the leading DNA-based techniques must be performed at the elevated temperature. However, our method does not require careful control of temperature by generating an isothermic enzymatic amplification of a visual signal. A color change from yellow to red on the simple surface in this study does not require sophisticated instrumentation unlike PCR requiring a thermal cycler. With regard to a reaction time, PCR could take upto a couple of hours whereas our system generates signal within 5-10 min. PCR requires trained personnel to prepare a reaction mixture including a DNA taq polymerase, primers, deoxyribonucleotide triphosphates (dNTPs), and a

reaction buffer. However, our method does not require these reagents such as carefully designed primers and a temperature-sensitive DNA polymerase enzyme by simply adding DNA and the detection ZFP on the surface.

In principal, a solution-based assay would allow more flexible orientations of ZFPs to fully interact with DNA as compared to immobilization of ZFPs on the 2D COC surface. However, immobilizing ZFPs on the COC surface allows us to wash off unbound molecules that could not be done in a solution-based assay. Also, ZFP immobilization can be a superior platform for developing a point-of-care detection device with the ability to simultaneously detect multiple pathogenic DNA sequences.[17] In our future study, a different immobilization method can be developed, which allows flexible orientations of ZFPs to capture more of target DNA, thus improving the sensitivity. In addition, we would also investigate a more sensitive detection method such a way to generate exponential signal rather than linear amplification of signal.

Specificity of engineered ZFP recognition

ZFP arrays were examined to demonstrate the specificity of the ZFP pairs in the presence of their own target DNA as well as non-target DNAs. Target site 1, 2, and 3 are the target DNA sequences for stx2_268/stx2_233, stx2_560/stx2_525, and stx2_1093/1128, respectively. An irrelevant DNA sequence being the target sites for Zif268 and PBSII was also included, which is not present in *E. coli* genome. As shown in Fig. 3, when a ZFP pair was incubated with its own target DNA, signal was distinctively high as compared to those of non-target and irrelevant DNAs. Thus, all of the ZFP pairs were able to distinguish their own targets from non-targets and irrelevant sequences. With high specificity, our ZFP array was able to detect specific double-stranded pathogenic DNA sequence. However, it is unsure if this level of differentiation

between target DNA and non-cognate DNA would be sufficient enough for real-world diagnostics, since we are still at the early stage of developing and optimizing our system with ZFPs, including further development of a new surface for better/optimized ZFP immobilization leading to capturing more DNA and an improved detection method with exponential signal. Since our assay generates a color change from yellow to red upon ZFP binding to its target DNA, the color change should be able to tell us whether or not the target DNA is detected as shown in digital image of nitrocefin assay in Fig. 3(A). In a real-world diagnostic that we envision in the future, the red color would indicate infection, the presence of a specific pathogen, and yellow color would imply no infection.

Understanding interactions between DNA and proteins has been one of the major issues to be addressed in biology and biophysics.[26] Understanding DNA-protein interactions provides us with a better understanding of recognition of their DNA targets and gene regulation. The *in vitro* binding motifs of our engineered ZFPs were determined by using Bind-n-Seq, a target site selection assay using massively parallel sequencing technology.[18] Interestingly, the motifs for stx2_233, stx2_268, and stx2_525 are found to be toward the middle, 5' end, and 3' end of their target sequences, respectively (Table 2). Since ZFPs have a sequence preference toward 5'-GNN-3',[13, 14, 27] their binding motifs may have been affected by the positions of ZF domains targeting 5'-GNN-3' triplet. It could also indicate that those motifs are the highest binding affinity binding regions of these proteins, and the regions that are primarily determining the specificity. The strong motifs were revealed from the ZFPs with high affinities. Likewise, the reasonable binding motif was not obtained for stx2_560 with low affinity (Supplementary data, Table S1). This data suggests that binding affinity obtained by EMSA may correlate with the identification of reasonable binding motifs.

DNA binding affinity

The binding specificity of ZFPs and their binding affinity toward their target DNAs are one of the key factors controlling the function of ZFP *in vivo*.[27] There might be multiple factors that affect the binding affinity of engineered ZFPs, demonstrating a wide range of affinities and specificities to their target sequences.[14, 15, 17, 27, 28] We examined the binding affinities of our engineered ZFPs toward their respective target DNA sequences using EMSA. Binding affinities of the ZFPs to their targets were determined to be in a range from 1.98 nM to 200 nM (Table 1 and Supplementary data), which are in the range of reported k_D values for six-finger ZFPs.[1, 14, 17] The k_D values of the capture probe ZFPs are 1.98, 200, and 75 nM for stx2_268, stx2_560, and stx2_1093, respectively, differing approximately by up to a 100-fold. The highest affinity capture probe was paired with the detection probe stx2_233 whose affinity is similar to that of the capture probe. The sensitivity of the stx2_268/stx2_233 pair differs by a 5-fold compared to those of the other two pairs. The combination of two very high-affinity ZFPs resulted in the highest sensitivity among the three ZFP pairs, without compromising assay specificity. The higher affinity ZFP pair may have retained its DNA for a longer period than the lower affinity pairs. The difference in binding affinity may have contributed to the differences in the sensitivity of a ZFP pair on the array.

Conclusions

We have demonstrated the direct detection of pathogen-specific dsDNA sequence utilizing engineered ZFPs arrayed on the copolymer chip. Our system avoids multiple laborious steps involved in DNA denaturation and subsequent hybridization, and DNA-labeling, providing a key

to a simple and rapid DNA sensing technology. Our ZFP pairs showed high specificity and sensitivity toward their own target DNA, suggesting that ZFP arrays on the COC chip could be further developed into a novel and reliable molecular device for pathogen detection. Our future study will focus on exploring immobilization methods for ZFPs and methods of signal amplification to achieve improved sensitivity. In near future, we envision a lab-on-a chip diagnostic by integrating our system into a microfluidic module. If more complex biological sample such as bacterial cell lysates is used on the microfluidic module integrated system, it would allow for pre-concentration of cell lysates.[29] Thus, it could lead us to lower the current limit of detection. In our previous study,[17] our assay system was still able to detect the target DNA in the presence of complex genomic DNA that would act as a competitor for specific binding. It was noted that ZFPs are stable at room temperature while the assay is being performed for many hours. Also, we would expect that ZFPs are fairly tolerant to the biological samples such as cell lysates, thus ZFP performance would not be affected by the cell lysate sample. Taken together, the stability and performance of ZFPs would not be a significant concern when using the biological sample.

In summary, our approach has demonstrated three different sets of ZFPs that specifically bind to three different target DNA sequences within the stx2 gene. For multiplexed detection, multiple ZFPs that recognize different target DNAs can be engineered and arrayed on the surface. The use of ZFPs makes our system novel because it does not require DNA denaturation and subsequent hybridization unlike the leading DNA-based methods. Our system generating a visual color change is suitable for POC diagnostics since it does not require labeling or sophisticated instrumentations in addition to the thermoplastic polymer COC surface as an ideal lab-on-a-chip platform. While ZFPs provide a powerful scaffold for custom-designed DNA-

binding proteins, a new class of DNA-binding domains TALEs (transcriptional activator-like effectors) could be engineered for a new diagnostic probe for detecting multiple pathogens. Compared to ZFPs, TALEs exhibit more modular architecture and flexibility for design, which could make TALE performance favorable for diagnostic application.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

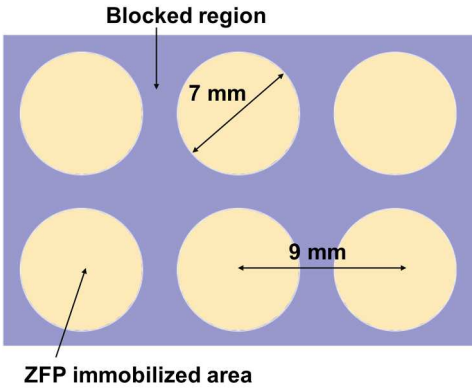
This research was supported by the Kentucky Biomedical Research Infrastructure Network (KBRIN) Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number 2P20GM103436-14 and National Science Foundation-International Research Experiences for Students (IRES) Award IIA-1358222.

References

1. M.-S. Kim and J. Kim, *Anal. Methods*, 2016, **8**, 6696-6700.
2. I. Ghosh, C. I. Stains, A. T. Ooi and D. J. Segal, *Mol. Biosyst.*, 2006, **2**, 551-560.
3. F. J. Pollock, P. J. Morris, B. L. Willis and D. G. Bourne, *Appl. Environ. Microbiol.*, 2010, **76**, 5282-5286.
4. T. Strachan and A. P. Read, *Human molecular genetics 3*, Garland Press, London ; New York, 3rd ed edn., 2004.
5. I. M. Mackay, K. E. Arden and A. Nitsche, *Nucleic Acids Res.*, 2002, **30**, 1292-1305.
6. R. Bumgarner, *Curr. Protoc. Mol. Biol.*, 2013, **Chapter 22**, Unit 22.21.
7. K. Zwirglmaier, *FEMS Microbiol. Lett.*, 2005, **246**, 151-158.
8. S. A. Wolfe, L. Nekludova and C. O. Pabo, *Annu. Rev. Biophys. Biomol. Struct.*, 2000, **29**, 183-212.

9. A. T. Ooi, C. I. Stains, I. Ghosh and D. J. Segal, *Biochemistry*, 2006, **45**, 3620-3625.
10. C. O. Pabo, E. Peisach and R. A. Grant, *Annu. Rev. Biochem.*, 2001, **70**, 313-340.
11. M. S. Kim and A. G. Kini, *Mol. Cells*, 2017, **40**, 533-541.
12. D. J. Segal, R. R. Beerli, P. Blancafort, B. Dreier, K. Effertz, A. Huber, B. Koksch, C. V. Lund, L. Magnenat, D. Valente and C. F. Barbas, 3rd, *Biochemistry*, 2003, **42**, 2137-2148.
13. B. Dreier, D. J. Segal and C. F. Barbas, 3rd, *J. Mol. Biol.*, 2000, **303**, 489-502.
14. D. J. Segal, B. Dreier, R. R. Beerli and C. F. Barbas, 3rd, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 2758-2763.
15. M. S. Bhakta and D. J. Segal, *Methods Mol. Biol.*, 2010, **649**, 3-30.
16. Q. Liu, D. J. Segal, J. B. Ghiara and C. F. Barbas, 3rd, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 5525-5530.
17. M. S. Kim, G. Stybayeva, J. Y. Lee, A. Revzin and D. J. Segal, *Nucleic Acids Res.*, 2011, **39**, e29.
18. A. Zykovich, I. Korf and D. J. Segal, *Nucleic Acids Res.*, 2009, **37**, e151.
19. S. H. Lee, S. W. Kim, J. Y. Kang and C. H. Ahn, *Lab on a chip*, 2008, **8**, 2121-2127.
20. R. Liedert, L. K. Amundsen, A. Hokkanen, M. Maki, A. Aittakorpi, M. Pakanen, J. R. Scherer, R. A. Mathies, M. Kurkinen, S. Uusitalo, L. Hakalahti, T. K. Nevanen, H. Siitari and H. Soderlund, *Lab on a chip*, 2012, **12**, 333-339.
21. S. Laib and B. D. MacCraith, *Anal. Chem.*, 2007, **79**, 6264-6270.
22. H. Becker and C. Gartner, *Electrophoresis*, 2000, **21**, 12-26.
23. C. H. Ahn, C. Jin-Woo, G. Beaucage, J. H. Nevin, L. Jeong-Bong, A. Puntambekar and J. Y. Lee, *Proceedings of the IEEE*, 2004, **92**, 154-173.
24. S. Kawasaki, P. M. Fratamico, N. Horikoshi, Y. Okada, K. Takeshita, T. Sameshima and S. Kawamoto, *Foodborne Pathog. Dis.*, 2010, **7**, 549-554.
25. B. Suo, Y. He, G. Paoli, A. Gehring, S. I. Tu and X. Shi, *Mol. Cell. Probes*, 2010, **24**, 77-86.
26. S. E. Halford and J. F. Marko, *Nucleic Acids Res.*, 2004, **32**, 3040-3052.
27. Q. Liu, Z. Xia, X. Zhong and C. C. Case, *J. Biol. Chem.*, 2002, **277**, 3850-3856.
28. J. D. Sander, P. Zaback, J. K. Joung, D. F. Voytas and D. Dobbs, *Nucleic Acids Res.*, 2009, **37**, 506-515.
29. I. F. Cheng, H. C. Chang, D. Hou and H. C. Chang, *Biomicrofluidics*, 2007, **1**, 21503.

(A)



(B)

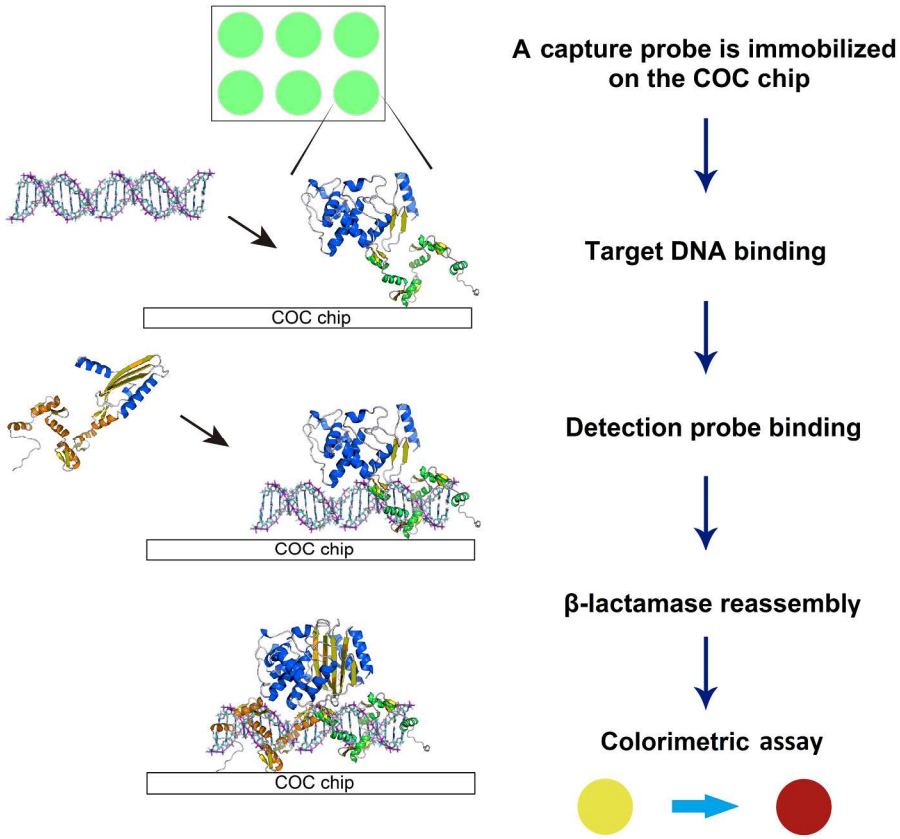
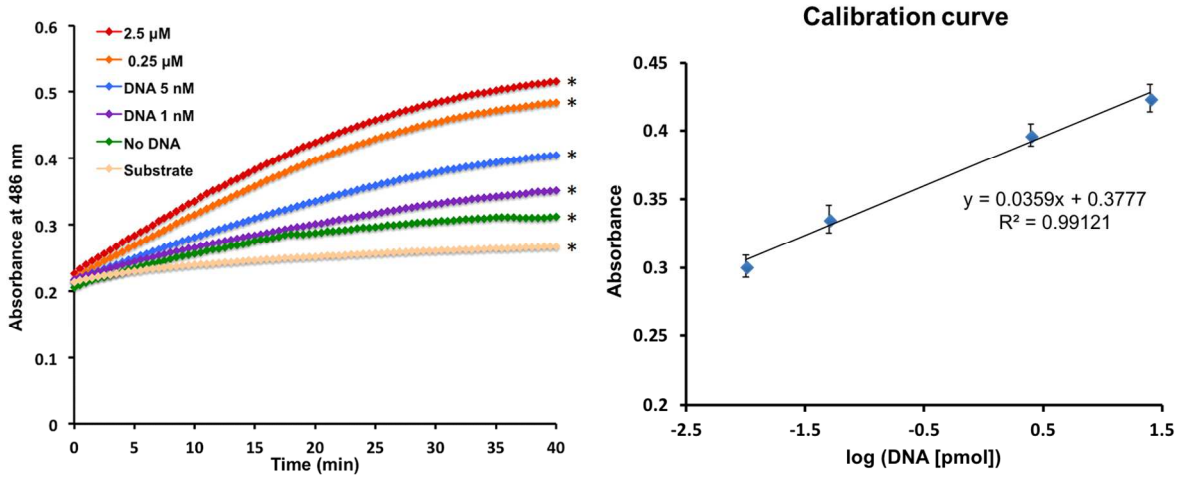
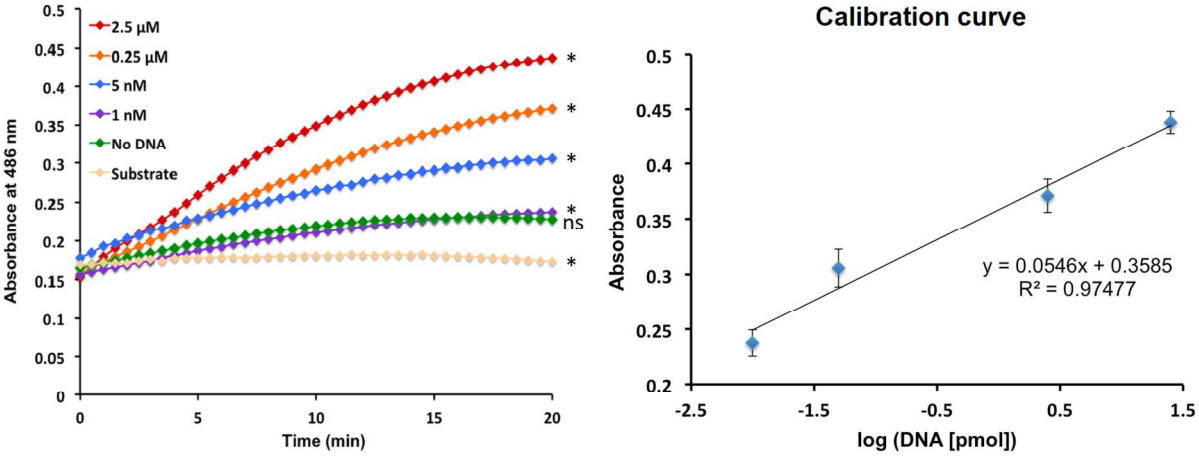


Fig. 1 (A) An image of COC array spot chip and (B) a schematic diagram of a ZFP array on the COC chip.

(A)


2.5 µM
0.25 µM
DNA 5 nM
DNA 1 nM
No DNA
Substrate
Absorbance at 486 nm
Time (min)
Calibration curve
y = 0.0359x + 0.3777
R² = 0.99121
Absorbance
log (DNA [pmol])

(B)


2.5 µM
0.25 µM
5 nM
1 nM
No DNA
Substrate
Absorbance at 486 nm
Time (min)
Calibration curve
y = 0.0546x + 0.3585
R² = 0.97477
Absorbance
log (DNA [pmol])

(C)

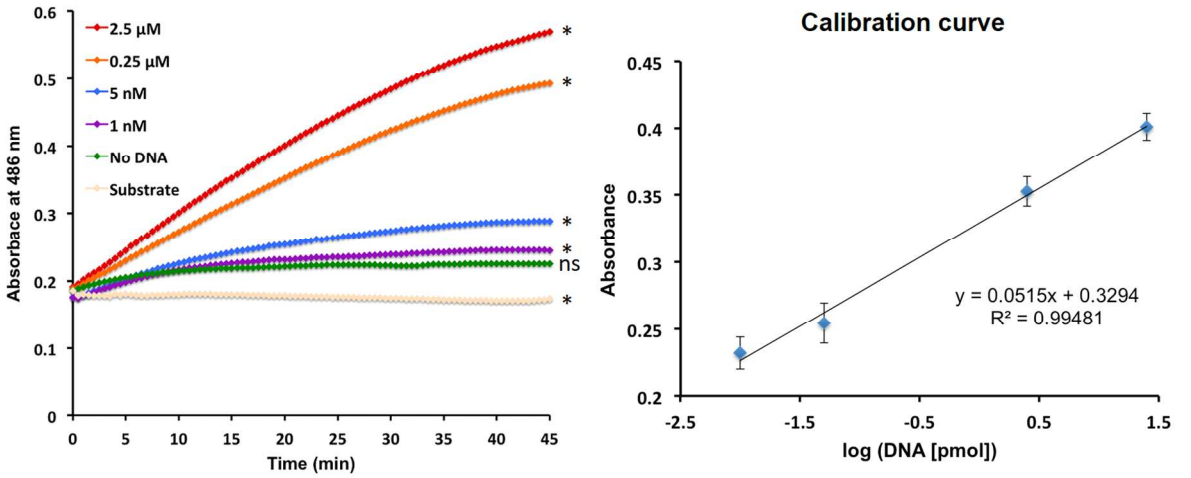
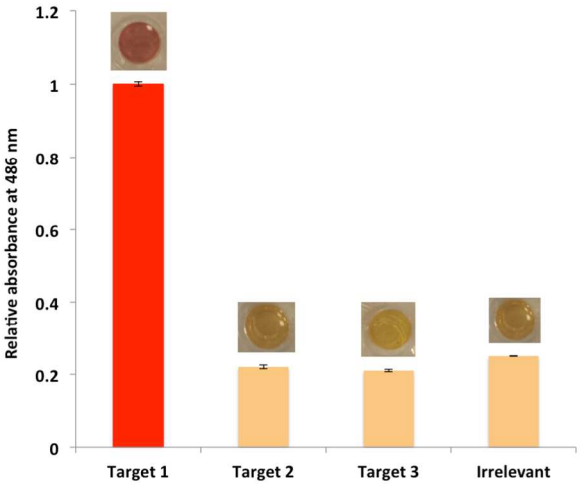
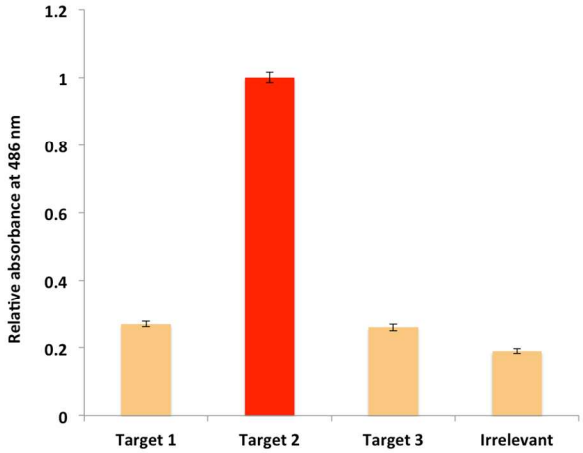

2.5 µM
0.25 µM
5 nM
1 nM
No DNA
Substrate
Absorbace at 486 nm
Time (min)
Calibration curve
y = 0.0515x + 0.3294
R² = 0.99481
Absorbance
log (DNA [pmol])

Fig. 2 The limit of detection of the three ZFP pairs. (A) The ZFP pair LacA stx2_268 and stx2_233 LacB, (B) the ZFP pair LacA stx2_560 and stx2_525 LacB, (C) the ZFP pair LacA stx2_1093 and stx2_1128 LacB. Final data points obtained after incubation with an asterisk (*) indicate significant differences ($P < 0.05$) (ns: not significant).

(A)


1.2
1
0.8
0.6
0.4
0.2
0
Relative absorbance at 486 nm
Target 1
Target 2
Target 3
Irrelevant

(B)


1.2
1
0.8
0.6
0.4
0.2
0
Relative absorbance at 486 nm
Target 1
Target 2
Target 3
Irrelevant

(C)

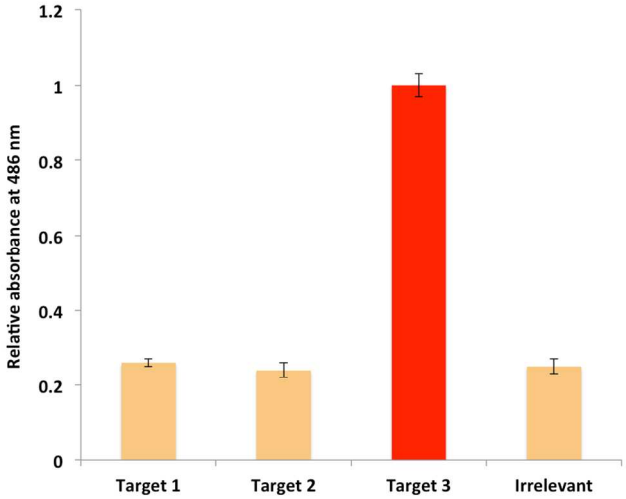

1.2
1
0.8
0.6
0.4
0.2
0
Relative absorbance at 486 nm
Target 1
Target 2
Target 3
Irrelevant

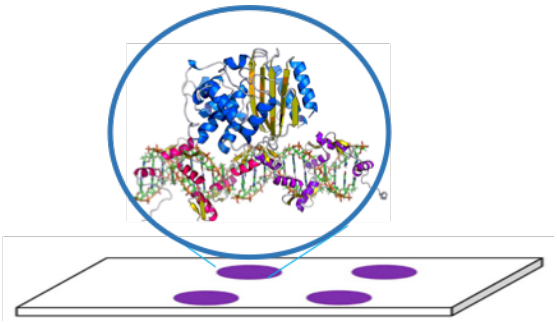
Fig. 3 The specificity of ZFP pairs. (A) The ZFP pair LacA stx2_268 and stx2_233 LacB, (B) the ZFP pair LacA stx2_560 and stx2_525 LacB, (C) the ZFP pair LacA stx2_1093 and stx2_1128 LacB. The corresponding digital image of nitrocefin assay is represented on the top of each bar graph in (A).

Table 1. Sequences of zinc finger recognition modules and their corresponding 3 bp DNA subsites, and the k_D values of zinc finger proteins.

ZFP	Position	Finger 6	Finger 5	Finger 4	Finger 3	Finger 2	Finger 1	k_D (nM)
Target site	233	CCT	GAT	AGA	CAT	CAA	GCC	3.56
Stx2_233		TSGNLVR	QSGNLTE	TSGNLTE	QLAHLRA	TSGNLVR	TKNSLTE	
Target site	268	GAA	GAT	GGT	CAA	AAC	GCG	1.98
Stx2_268		RSDDLVR	DSGNLRV	QSGNLTE	TSGHLVR	TSGNLVR	QSSNLVR	
Target site	525	AGA	ACT	GCT	CTG	GAT	GCA	4.50
Stx2_525		QSGDLRR	TSGNLVR	RNDALTE	TSGELVR	THLDLIR	QLAHLRA	
Target site	560	GTG	ACA	GTG	ACA	AAA	CGC	200
Stx2_560		HTGHLLE	QRANLRA	SPADLTR	RSDELVR	SPADLTR	RSDELVR	
Target site	1093	AAG	TGG	CCG	GGA	AAG	AAT	75
Stx2_1093		TTGNLTV	RKDNLN	QRAHLE	RNDTLTE	RSDHLTT	RKDNLKN	
Target site	1128	ACT	GGA	CCA	GTC	GCT	GGA	51.3
Stx2_1128		QRAHLER	TSGELVR	DPGALVR	TSHSLTE	QRAHLER	THLDLIR	

Table 2. Binding motifs identified by the Bind-n-Seq for stx2_233, stx2_268, and stx2_525.

	Motifs	Seed	Enrichment	Protein	Salt
Stx2_233	CCTGATAGACATCAAGCC 	AAGATAGACC	42.667	50 nM	100 mM
Stx2_268	GAAGATGGTCAAAACGCG 	AGGAAGATGG	86.750	5 nM	100 mM
Stx2_525	AGAACTGCTCTGGATGCA 	ATGGATGCA	115.500	50 mM	50 mM



An array of engineered Zinc Finger Proteins functions as sensing pathogen-specific DNA along with detection probes on cyclic olefin copolymer chip.