

An Improved Method for Identifying Specific DNA-Protein-Binding Sites In Vitro

Liangyan Wang¹ · Huizhi Lu¹ · Yunguang Wang¹ · Su Yang¹ · Hong Xu¹ · Kaiying Cheng¹ · Ye Zhao¹ · Bing Tian¹ · Yuejin Hua¹

Published online: 28 January 2017
© Springer Science+Business Media New York 2017

Abstract Binding of proteins to specific DNA sequences is essential for a variety of cellular processes such as DNA replication, transcription and responses to external stimuli. Chromatin immunoprecipitation is widely used for determining intracellular DNA fragments bound by a specific protein. However, the subsequent specific or accurate DNA-protein-binding sequence is usually determined by DNA footprinting. Here, we report an alternative method for identifying specific sites of DNA-protein-binding (designated SSDP) in vitro. This technique is mainly dependent on antibody-antigen immunity, simple and convenient, while radioactive isotope labeling and optimization of partial degradation by deoxyribonuclease (DNase) are avoided. As an example, the specific binding sequence of a target promoter by DdrO (a DNA damage response protein from *Deinococcus radiodurans*) in vitro was determined by the developed method. The central sequence of the binding site could be easily located using this technique.

Keywords Chromatin immunoprecipitation · DNA footprinting · DNA-protein binding · Specific DNA-binding site · Antibody-antigen immunity

Abbreviations

ChIP	Chromatin immunoprecipitation
SSDP	Specific sites of DNA-protein-binding
DdrO	A DNA damage response protein from <i>Deinococcus radiodurans</i>
EMSA	Electrophoretic mobility shift assay
SPR	Surface plasmon resonance analyses
PBM	Protein-binding microarray
ChIA-PET	Chromatin interaction analysis by paired-end tag sequencing
DHS	DNase I hypersensitive sites
ENCODE	Encyclopedia of DNA elements
DDR	DNA damage response
<i>recA</i>	<i>Recombinant A</i>
MBP	Maltose-binding protein
TEV	Tobacco etch virus
IPTG	Isopropyl β-D-1-thiogalactopyranoside
BPB	Bromophenol blue
RDRM	Radiation and desiccation response motif

Liangyan Wang and Huizhi Lu have contributed equally to the paper as first authors.

✉ Bing Tian
tianbing@zju.edu.cn

✉ Yuejin Hua
yjhua@zju.edu.cn

¹ Key Laboratory of Chinese Ministry of Agriculture and Zhejiang Province for Nuclear Agricultural Sciences, Institute of Nuclear Agricultural Sciences, Zhejiang University, No. 268, Kaixuan Road, Hangzhou 310029, Zhejiang, China

Introduction

DNA and proteins often interact to participate in a variety of biological processes, such as DNA replication, transcription and repair. For example, transcription factors always bind to particular sets of DNA sequences, activating or inhibiting the transcription of their downstream genes [1–4]. Consequently, alterations in the activity of any transcription factor can affect an array of target genes. Thus, the transcription factors could control diverse signal

transduction processes, responding to cellular differentiation, development and environmental changes. For decades, numerous approaches were developed for determining specific DNA-protein-binding in vivo or in vitro, including affinity chromatography [5], chromatin immunoprecipitation (ChIP) [6], electrophoretic mobility shift assays (EMSA) [7, 8], southwestern blotting [9], surface plasmon resonance analyses (SPR) [10] and protein-binding microarray (PBM) [11]. Among them, ChIP is widely used to determine the particular DNA fragments bound by a target protein of interest in vivo. In this method, DNA-protein complexes are cross-linked, sheared into ~500 bp DNA fragments by sonication and then immunoprecipitated with antibodies raised against the target protein. As a result, the associated DNA fragments in cellular context are enriched and their sequences could be determined [5, 6]. Afterward, chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) technology that incorporates chromatin immunoprecipitation (ChIP)-based enrichment, chromatin proximity ligation, paired-end tags and high-throughput sequencing was developed to identify unique and functional chromatin interactions between distal and proximal regulatory transcription factor-binding sites and the promoters of the genes [12].

Though ChIP has been an efficient way to determine in vivo specific protein-associated DNA fragments, accurate detection of protein-protected DNA sequence is usually dependent on the classical in vitro footprinting method. Currently, the DNase I footprinting assay has been extensively utilized for both identification and characterization of specific DNA-protein-binding site because of its simplicity and efficiency [13]. Binding of proteins to specific regions of the genome often leads to looseness of condensed chromatin, so as to produce DNase I hypersensitive sites (DHSs) which are sensitive to cleavage by the DNase I enzyme. The famous encyclopedia of DNA elements (ENCODE) project was aimed to map all of the DHSs in the human or other genomes to identify all regulatory DNA elements [14]. DNase I footprinting assay employs DNase I to partially digest the radioactively end-labeled DNA, generating a ladder of DNA fragments, followed by the use of denaturing acrylamide gel electrophoresis to detect the resulting cleavage pattern. Positions in a subsequent autoradiograph will represent the distance from the end label to the points of cleavage, which makes it possible to locate a protein-binding site on a particular DNA molecule. To acquire a clear “footprint”, both the quantity of DNase I and the minimum concentration of the DNA-binding protein should be optimized.

Recently, the product of an orphan gene in *Deinococcus*, DdrO, has been shown to play an important role in DNA damage response in *Deinococcus radiodurans* [15, 16].

Our previous results demonstrated that DdrO could directly bind to the promoter regions of the DNA damage response genes (DDR) including *recombinant A (recA)* in vitro. Here, we report a simple, convenient and radioactive isotope-free method for identifying specific DNA-protein-binding sites, designated SSDP (specific sites of DNA-protein-binding), which incorporate the superiority of both CHIP and DNA footprinting. As an example, the specific DNA-binding sequence of the target *recA* promoter by DdrO is tested, and the central DNA sequence of the binding site is readily located.

Materials and Methods

Protein Expression and Purification

The DdrO protein expression and purification were performed as previously described [16]. Briefly, the *ddro* gene was cloned into the pET28-HMT vector with a hexahistidine tag and a maltose-binding protein (MBP) tag separated by the tobacco etch virus (TEV) protease site. The recombinant product was then introduced into *Escherichia coli* (*E. coli*) BL21 (DE3) for subsequent expression. Cells were induced by isopropyl β -D-1-thiogalactopyranoside (IPTG) for 4 h, harvested by centrifugation (12,000 \times g) and disrupted in lysis buffer [250 mM NaCl, 20 mM Tris-HCl pH 8.0, 1 mg/ml lysozyme, 1000 U/ml DNase I, 1 mM phenylmethanesulfonylfluoride (PMSF)]. The supernatant was loaded onto a Ni-NTA Superflow column using the AKTA Purifier System (GE Healthcare, Pittsburgh, USA). Protein was then separated on an amylose column, eluted and subsequently cleaved by TEV protease. The sample was further purified by Ni-NTA and MBP column chromatography to discard the 6 \times His-MBP tag. To obtain highly pure protein, sequential HiTrap Heparin HP and gel filtration chromatography (Superdex 75) were used.

Affinity Purification of Polyclonal Antibody

Affinity purification of antibody was performed as previously described with some modifications [17]. Following purification of IgG using protein A/G agarose (Life Technologies, Carlsbad, California, USA), the antigen-specific antibody purification was performed. To fill the gravity column, 10 ml of activated amino-coupled agarose resin was taken out. Then, 8 ml of DdrO protein (0.5 mg/ml) and 125 μ l of 5 M CH_3BNNa (dissolved in 1 M NaOH) were loaded onto the 1 \times PBS-washed column and incubated for 6 h at room temperature while rotating. The un-cross-linked activated group was packaged by 1 M Tris-HCl (pH 7.4) and removed. Thus, the DdrO affinity

chromatography column was prepared. Afterward, DdrO antibodies (rabbit IgG, laboratory stock) were applied to the above-prepared resin and incubated for 2 h at room temperature. The specific antibodies were eluted with 0.1 M glycine-HCl (pH 2.5) and neutralized with 1 M Tris-HCl (pH 9.0). The quantity of the antibody was measured by the Coomassie brilliant blue method [18]. Samples were finally dialyzed in the storage buffer [200 mM NaCl, 20 mM Tris-HCl pH 8.0, 0.2 mM dithiothreitol (DTT)], concentrated and stored in small packages in liquid nitrogen.

Formaldehyde Chemical Cross-linking

Protein-protein chemical cross-linking for DdrO-DdrO was performed as previously described with some modifications [19]. Purified proteins were diluted to 1.7 mg/ml in 50 mM HEPES buffer (pH 8.2) with 0.1 mM EDTA. A final concentration of 25 mM formaldehyde solution was added to the 120 μ l reaction and incubated at 30 °C for 30 min. Equal volume of SDS buffer was then added to stop the reaction. Protein interactions were visualized by SDS-PAGE electrophoresis.

Preparation of Protein-Protective Oligonucleotide

After optimizing the ratio of the DNA fragment to protein, the promoter of *recA* gene (PCR products with primers *PrecA-F*: 5'-AGAAACACCAGCATGATCGC-3' and *PrecA-R*: 5'-CGTGCCCACACTGATGAT-3', 187 bp) was incubated with DdrO in 250 μ l of binding buffer (150 mM NaCl, 20 mM Tris-HCl, 5 mM MgCl₂) for 1 h at 30 °C. Five microliters of binding buffer was mixed with 5 \times electrophoretic mobility shift assay (EMSA) loading buffer (150 mM NaCl, 20 mM Tris-HCl, 5 mM MgCl₂, 0.2% (w/v) bromophenol blue (BPB), 37% glycerol) to detect DNA-protein-binding using the EMSA method. The remaining buffer was incubated with DNase I (100 U) and CaCl₂ (1 mM final concentration) for 40 min to digest the unbound DNA sequence. Then, 100 μ l of protein A/G agarose resin (Life Technologies) containing purified DdrO antibody was added and incubated for 3 h while rotating. Nonspecific anti-rabbit antibody was used as the negative control. Following being washed by binding buffer (150 mM NaCl, 20 mM Tris-HCl, 5 mM MgCl₂) for several times, the antibody-DdrO-DNA complex was eluted by elution buffer (1% SDS in 0.1 M NaHCO₃). After the unbound DNA was digested by DNase I, a final concentration of 0.5 M EDTA and 200 μ g/ml proteinase K was added to eliminate binding protein. The specific binding DNA fragments were finally extracted by the MinElute Gel Extraction Kit (Qiagen, Valencia, CA, USA).

Sequencing of the Specific Oligonucleotide

Both ends of oligonucleotide were filled and blunted with T4 DNA polymerase (NEB, Ipswich, MA, England), armed with the adapters (Adapter 1: 5'-AGCGAATAAGA GCTTGTCTCCAGGTACGCTAGCAATGCTCTCGAA CCTTAGCTGGTA-3' dsDNA; Adapter 2: 5'-AGGCGA GAGAATCATTATCAAGGAGGTCTGGGACAAGCCA TCCTAGAATCCGACTATGT-3' dsDNA) and then ligated to the pMD 18-T simple vector (TaKaRa, Dalian, China). The resultant recombinant vector was finally sequenced to determine the accurate sequence of inserted, specific DNA fragments with the sequencing primers of the pMD 18-T simple vector.

Results

Scheme for Detecting the Specific DNA-Protein-Binding Site

To eliminate the introduction of radioactive isotope and shorten the experimental period, we developed a relatively convenient and rapid method called SSDP to identify specific or accurate DNA-protein-binding sites in vitro. The essential steps of this procedure are shown in Fig. 1. Initially, this approach was based on highly specific biological interactions between antibody and antigen, so specific antibody with high purity was required. Meanwhile, extended promoter DNA fragments having been proved to be bound by the natural specific protein were amplified and purified. Next, the optimal binding system was established by adjusting the ratio of protein to DNA fragment so that maximum DNA-protein complexes were obtained. After the flanking DNA fragments were digested by DNase I, protein-protected DNA sequences were co-purified with the antibody mediated by the protein. Then, the collected specific DNA fragments were stripped from the binding proteins, sequentially ligated to specifically designed adapters and the T-vector, and finally sequenced.

Refinement of DdrO Antibody for Efficient Binding

Recently, we have verified that the product of an orphan gene *ddrO* in *D. radiodurans* could specifically bind to the *recA* promoter in vitro [16]. Here, we would like to reveal the specific or accurate *recA* promoter-DdrO-binding site with this improved method. The purity of antibody is of great importance for this method, so a two-step affinity purification process (protein A/G affinity followed by DdrO affinity) for refining polyclonal DdrO antibody was used. As shown in Fig. 2a, the purified antibodies displayed a more specific and reactive binding pattern, with

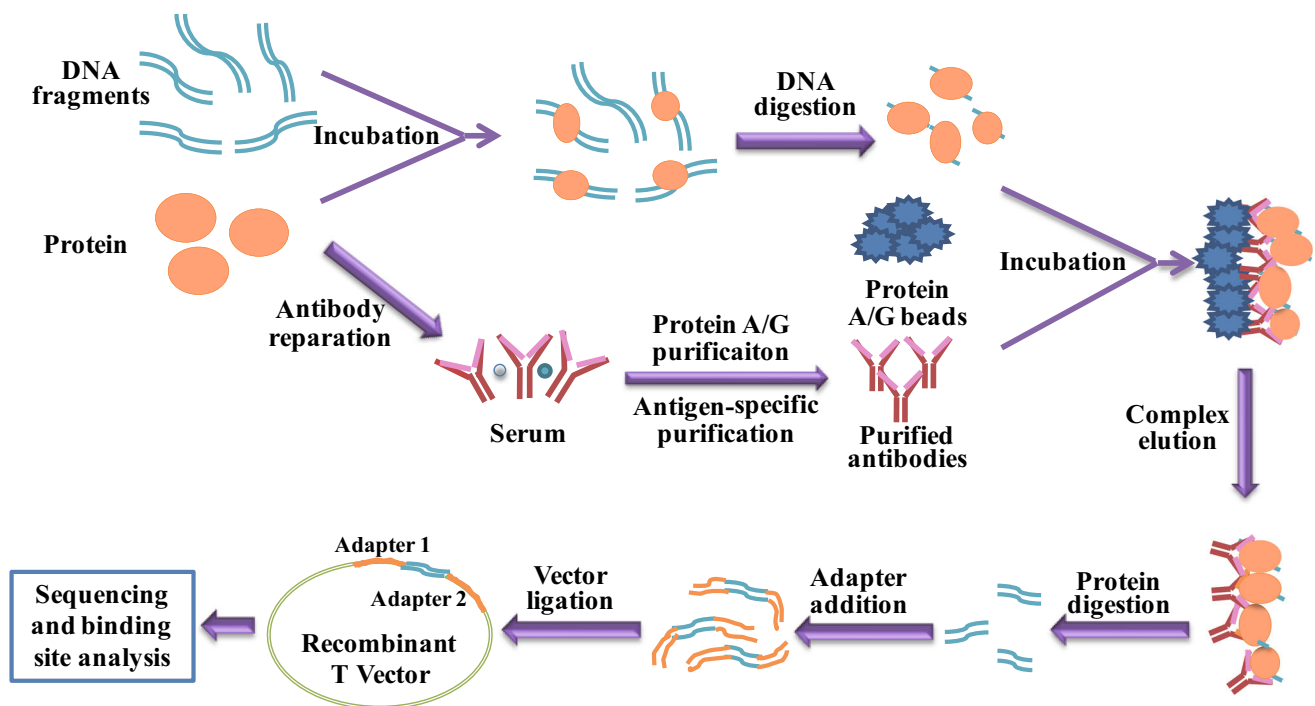


Fig. 1 Schematic diagram for the newly developed method of detecting specific DNA-protein-binding sites

more than twice the specificity relative to the original polyclonal antibodies (Fig. 2b). The highly purified polyclonals increased the yield of specific DdrO-binding DNA fragments, which is critical for this improved method.

Detection of DdrO Active Form for Binding System Optimization

DdrO is regarded as an essential transcription repressor in *D. radiodurans*, governing a series of DNA damage response genes [16]. To further elucidate its functional mechanism, the homopolymeric form of purified protein was determined by gel filtration and cross-linking method, respectively, in vitro. As shown in Fig. 2c, d, the results of both gel filtration assay and chemical cross-linking assay demonstrated that DdrO functions as a dimer. Because accurate quantities of protein and oligonucleotide are essential for the specific DNA-protein-binding system, this information will help us to confirm the activity of DdrO and determine the optimal ratio of DdrO binding to the promoter DNA fragment.

Determination of Specific Binding Site of DdrO-*recA* Promoter

To probe the accurate DNA-binding sequence of the target protein, current established method was utilized. The optimum ratio of DdrO to *recA* promoter (*PrecA*) in certain length was achieved through several trials using different

concentrations of DdrO and *recA* promoter. As indicated, when the concentration of *PrecA* was 0.8 μM , the optimum concentration of DdrO was approximately 8 μM (Fig. 3a). It is of importance to optimize the reaction system, assuring that all DNA fragments are fully associated with the protein, whereas the least amount of free protein is left. After free DNA was completely digested by DNase I, oligonucleotides protected by protein were isolated and purified. Then, the adapters (see “Materials and Methods”) were fused to both ends of oligonucleotides, respectively, followed by PCR confirmation. As shown in Fig. 3b, oligonucleotides armed with adapters at both ends displayed an expected band, while incomplete ligation resulted in nonspecific bands. The authentic target band was finally purified, amplified and ligated to the T-vector. Totally, ten individual plasmids are selected for sequencing. The results showed that all the sequences are “TTGTTATGCTGCTAGCAGAAATC” (Fig. 3c), about 3 base pairs (bp) longer than the predicted radiation and desiccation response motif (RDRM) site [16] at both ends. Combined with the previous EMSA result [16], the “3 bp” could be steric exclusion of the digestion enzyme by the bound protein.

Discussion

Identification of accurate or specific DNA-protein-binding sites in vitro has been an area of interest for gene transcription and other biological processes [20, 21]. As one

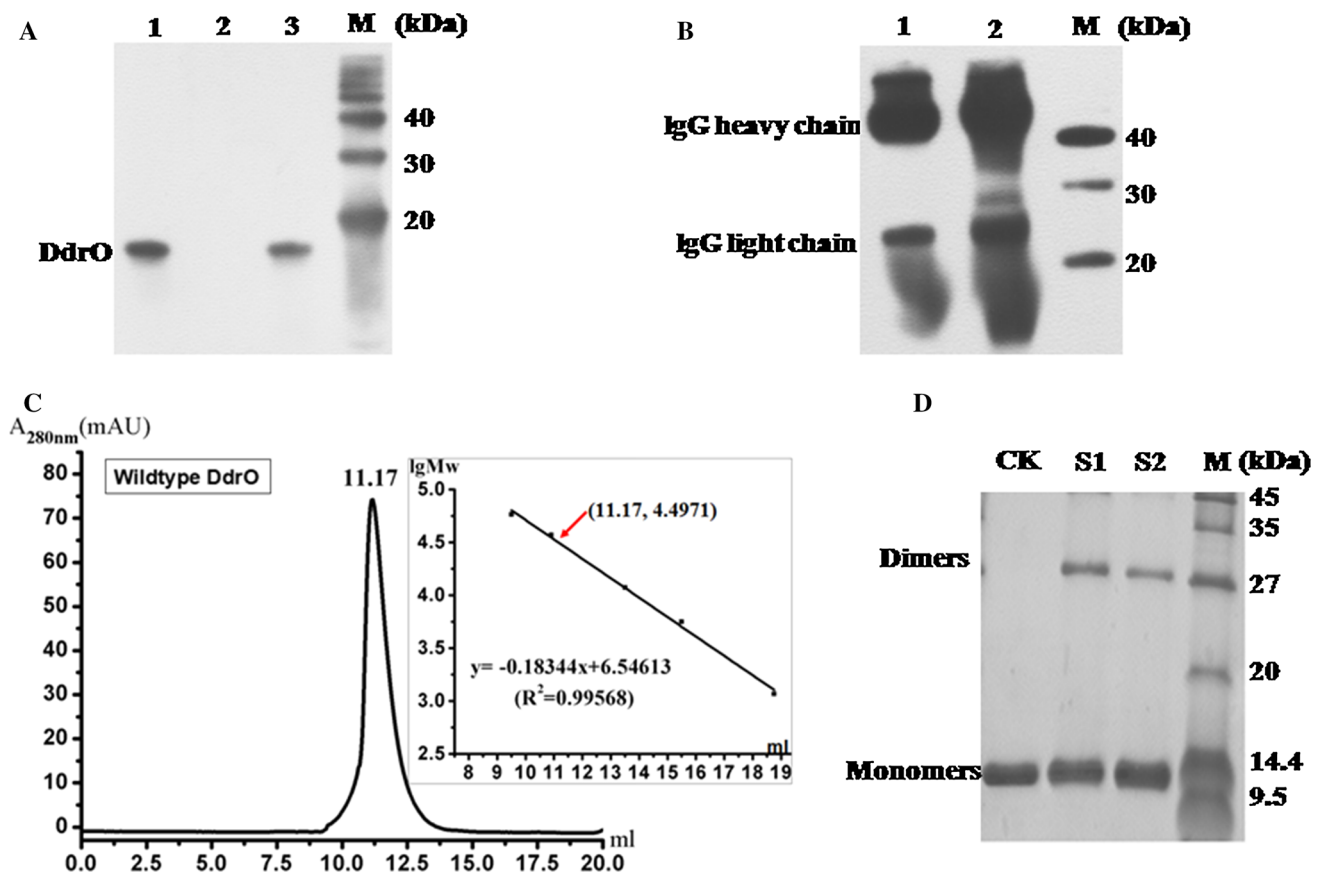


Fig. 2 Quality determination of DdrO polyclonal antibody. **a** Detection of DdrO with polyclonal antibody before and after affinity purification. *Lane 1* Antibody serum before purification; *Lane 2* antibody serum flow-through after purification; *Lane 3* antibody serum toward DdrO after purification; *Lane 4* protein markers. **b** Detection of DdrO polyclonal antibody with the secondary antibody. *Lane 1* Antibody serum before purification; *Lane 2* antibody serum after purification; *Lane 3* protein markers. **c** Detection of the molecular weight (MW) of native DdrO by gel filtration in vitro. Calibration of gel filtration column (Superdex 75, GE) was

based on proteins of known MW (BSA, MW 67,000; ovalbumin, MW 43,000; ribonuclease, MW 13,700; aprotinin, MW 6512; vitamin, MW 1355). The standard curve in the inserted figure was established by plotting retention volume date against the logarithm of the molecular weights of the known proteins. The molecular weight of DdrO dimer is approximately 31.4 kD. **d** Detection of the polymer form of DdrO using formaldehyde chemical cross-linking in vitro. *CK* indicates uncross-linked control samples; *S1* and *S2* both indicate cross-linked samples; *M* indicates protein markers

important and believable method, DNA footprinting has been extensively used, though potentially toxic and costly radioisotope is introduced.

In this study, we conceived an improved method based on specific antigen-antibody detection, using the specific DNA binding of the *recA* promoter by inhibitory transcription factor DdrO as a test system. Compared with other existing methods, this developed technology possesses several advantages. First, isotope labeling was eliminated to avoid health issues and reduce experimental cost. Second, DNA sequencing technology was used to analyze DNA-binding sites occupied by protein, providing more accurate results. Third, unlike DNA footprinting, DNase I digestion was complete, while no strict concentration of enzyme is needed; thus, the entire process was simple and easy to perform.

However, the current method is required to prepare polyclonal antibody with high purity. From the view of

purity, monoclonals are ideal choice, avoiding polyclonals purification via antigens. Nevertheless, polyclonals often bind much more antigens than equivalent monoclonals, while sufficient DNA fragments are beneficial for this method. So we still recommend to employ polyclonals with high purity. A successive two-step affinity purification process (protein A/G affinity and antigen affinity) could often meet the requirements. Alternatively, some universal tag antibodies such as anti-FLAG-M2 antibody may be used for detecting tagged protein. Meanwhile, undesirable recovery of protein-protective DNA fragments following DNase I digestion could lead to nonspecific background which might interfere with the subsequent results. And also, adapters are not always efficiently attached to the recycled DNA fragments, due to the presence of blunt ends and nondirectional adapters. Even so, adequate oligonucleotides in high quality could be easily obtained for

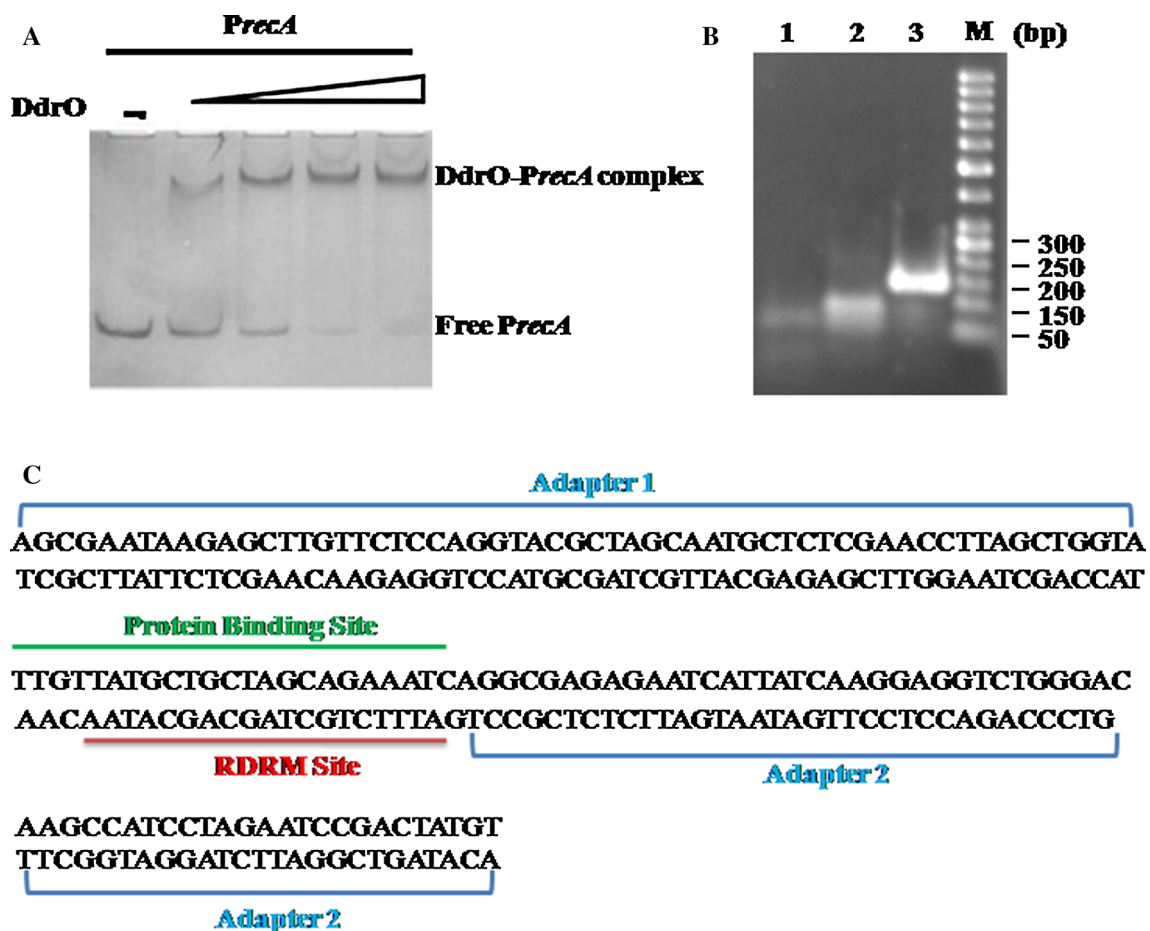


Fig. 3 Optimization of the protein–DNA-binding system and analysis of oligonucleotide adapter ligation. **a** The optimal ratio of DdrO to DNA fragments. The concentration of *recA* promoter (*Preca*) was 0.8 μM , while the concentration of DdrO was 0, 3, 4, 6 and 8 μM , respectively. **b** Confirmation of oligonucleotides ligated to the adapters. *Lane 1* Negative control containing only purified DNA fragments; *Lane 2* negative control containing only adapters; *Lane 3* PCR products from adapters and oligonucleotide ligation reaction;

M DNA marker 50 base pairs (MBI). **c** The specific DdrO-*recA* promoter-binding site. The sequence marked with a *green line* indicates the DdrO-*recA*-binding site identified by the current method. The sequence underlined in *red* indicates the predicted conserved sequence of radiation and desiccation response motif (RDRM) site. The flanking sequences indicate the adapters employed in this experiment (Color figure online)

subsequent processing through optimization of reaction systems. Hence, this method provides a feasible way to determine the relatively specific or accurate binding sites of DNA-protein (or other small molecules as well) complexes [22].

About a decade ago, the improved protein-binding microarray (PBM) technology was used to characterize transcription factors' sequence specificities in a high-throughput manner [11]. It has been a true, quick, unbiased and efficient tool for specific protein-DNA interactions in vitro. In comparison, the current technique is merely a relatively cheap, simple and efficient tool for the detection of individual protein-DNA interactions so far. Nevertheless, subsequent efforts could be made to identify DNA-binding profiles, for instance, determining all one-base variations in a transcription factor-DNA-binding profile.

Acknowledgement This work was supported by grants from the National Natural Science Foundation of China (31210103904, 31370102, 31570058), the Natural Science Foundation of Zhejiang Province (LY13C010001) and the Public Project of Zhejiang Province (2014C33024) and the Project for Experimental Technology of Zhejiang University.

Compliance with Ethical Standards

Conflict of interest The authors declare no financial or commercial conflict of interest.

References

- Liu, C., Wang, L., Li, T., Lin, L., Dai, S., Tian, B., et al. (2014). A PerR-like protein involved in response to oxidative stress in the extreme bacterium *Deinococcus radiodurans*. *Biochemical and Biophysical Research Communications*, 450, 575–580.

2. Yamazaki, S., Hayano, M., & Masai, H. (2013). Replication timing regulation of eukaryotic replicons: Rif1 as a global regulator of replication timing. *Trends in Genetics*, *29*, 449–460.
3. Nasmith, C. G., Walkowiak, S., Wang, L., Leung, W. W., Gong, Y., Johnston, A., et al. (2011). Tri6 is a global transcription regulator in the phytopathogen *Fusarium graminearum*. *PLoS Pathogens*, *7*, e1002266.
4. Wang, L., Xu, G., Chen, H., Zhao, Y., Xu, N., Tian, B., et al. (2008). DrRRA: A novel response regulator essential for the extreme radioresistance of *Deinococcus radiodurans*. *Molecular Microbiology*, *67*, 1211–1222.
5. Merante, F., Altamentova, S. M., Mickle, D. A., Weisel, R. D., Thatcher, B. J., Martin, B. M., et al. (2002). The characterization and purification of a human transcription factor modulating the glutathione peroxidase gene in response to oxygen tension. *Molecular and Cellular Biochemistry*, *229*, 73–83.
6. Mitra, S., Rai, L. S., Chatterjee, G., & Sanyal, K. (2016). Chromatin immunoprecipitation (ChIP) assay in *Candida albicans*. *Methods in Molecular Biology*, *1356*, 43–57.
7. Deckmann, K., Rorsch, F., Geisslinger, G., & Grosch, S. (2012). Identification of DNA–protein complexes using an improved, combined western blotting–electrophoretic mobility shift assay (WEMSA) with a fluorescence imaging system. *Molecular BioSystems*, *8*, 1389–1395.
8. Ruscher, K., Reuter, M., Kupper, D., Trendelenburg, G., Dirnagl, U., & Meisel, A. (2000). A fluorescence based non-radioactive electrophoretic mobility shift assay. *Journal of Biotechnology*, *78*, 163–170.
9. Labbe, S., Harrisson, J. F., & Seguin, C. (2009). Identification of sequence-specific DNA-binding proteins by southwestern blotting. *Methods in Molecular Biology*, *543*, 151–161.
10. Stockley, P. G., & Persson, B. (2009). Surface plasmon resonance assays of DNA–protein interactions. *Methods in Molecular Biology*, *543*, 653–669.
11. Berger, M. F., & Bulyk, M. L. (2009). Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. *Nature Protocols*, *4*, 393–411.
12. Fullwood, M. J., & Ruan, Y. (2009). ChIP-based methods for the identification of long-range chromatin interactions. *Journal of Cellular Biochemistry*, *107*, 30–39.
13. Galas, D. J., & Schmitz, A. (1978). DNase footprinting: A simple method for the detection of protein–DNA binding specificity. *Nucleic Acids Research*, *5*, 3157–3170.
14. Sloan, C. A., Chan, E. T., Davidson, J. M., Malladi, V. S., Strattan, J. S., Hitz, B. C., et al. (2016). ENCODE data at the ENCODE portal. *Nucleic Acids Research*, *44*, D726–D732.
15. Devigne, A., Ithurbide, S., Bouthier de la Tour, C., Passot, F., Mathieu, M., Sommer, S., et al. (2015). DdrO is an essential protein that regulates the radiation desiccation response and the apoptotic-like cell death in the radioresistant *Deinococcus radiodurans* bacterium. *Molecular Microbiology*, *96*, 1069–1084.
16. Wang, Y., Xu, Q., Lu, H., Lin, L., Wang, L., Xu, H., et al. (2015). Protease activity of PprI facilitates DNA damage response: Mn(2+)–dependence and substrate sequence-specificity of the proteolytic reaction. *PLoS ONE*, *10*, e0122071.
17. Arora, S., Ayyar, B. V., & O’Kennedy, R. (2014). Affinity chromatography for antibody purification. *Methods in Molecular Biology*, *1129*, 497–516.
18. Pierce, J., & Suelter, C. H. (1977). An evaluation of the Coomassie brilliant blue G-250 dye-binding method for quantitative protein determination. *Analytical Biochemistry*, *81*, 478–480.
19. Tang, X., & Bruce, J. E. (2009). Chemical cross-linking for protein–protein interaction studies. *Methods in Molecular Biology*, *492*, 283–293.
20. Kagawa, W., Kagawa, A., Saito, K., Ikawa, S., Shibata, T., Kurumizaka, H., et al. (2008). Identification of a second DNA binding site in the human Rad52 protein. *Journal of Biological Chemistry*, *283*, 24264–24273.
21. Chan, T. M., Leung, K. S., Lee, K. H., Wong, M. H., Lau, T. C., & Tsui, S. K. (2012). Subtypes of associated protein–DNA (transcription factor–transcription factor binding site) patterns. *Nucleic Acids Research*, *40*, 9392–9403.
22. Bailly, C., Kluza, J., Martin, C., Ellis, T., & Waring, M. J. (2005). DNase I footprinting of small molecule binding sites on DNA. *Methods in Molecular Biology*, *288*, 319–342.