Published in final edited form as:

Anal Biochem. 2011 August 15; 415(2): 168–174. doi:10.1016/j.ab.2011.03.027.

A filter microplate assay for quantitative analysis of DNA-binding proteins using fluorescent DNA

William C. Yang¹ and James R. Swartz^{1,2}

¹Department of Bioengineering, Stanford University, Stanford, CA 94305

²Department of Chemical Engineering, Stanford University, Stanford, CA 94305

Abstract

We present a rapid method for quantifying the apparent DNA binding affinity and capacity of recombinant transcription factors (TFs). We capture His₆-tagged TFs using Ni-NTA agarose and incubate the immobilized TFs with fluorescently-labeled cognate DNA probes. After washing, the strength of the fluorescence signal indicates the extent of DNA binding. The assay was validated using two pluripotency-regulating TFs, SOX2 and NANOG. Using competitive binding analysis with non-labeled competitor DNA, we show that SOX2 and NANOG specifically bind to their consensus sequences. We also determined the apparent affinity of SOX2 and NANOG for their consensus sequences to be 54.2±9 nM and 44.0±6 nM, respectively, in approximated agreement with literature values. Our assay does not require radioactivity, but radioactively labeling the TFs enables the measurement of absolute amounts of immobilized SOX2 and NANOG, and hence a DNA to protein binding ratio. SOX2 possesses a 0.95 DNA to protein ratio while NANOG possesses a 0.44 ratio, suggesting that most of the SOX2 and approximately half of the NANOG are competent for DNA binding. Alternatively, the NANOG dimer may be capable of binding only one DNA target. This flexible DNA binding assay enables the analysis of crude or purified samples with or without radioactivity.

Keywords

Cell-free protein synthesis; transcription factor; DNA binding

Introduction

In recent years, it has been shown that ectopic viral expression of transcription factors (TFs) can reprogram somatic cells to a pluripotent state [1; 2], control cell fate [3], and directly convert somatic cells from one lineage to that of another [4]. These results indicate the possibility of patient-specific cell therapies, but the use of viruses raises safety concerns. Viruses integrate exogenous DNA into the host cell genome and can disrupt crucial genes and/or induce tumorigenicity [5]. Therefore, in order to safely reprogram cells and bring the technology one step closer to the clinic, researchers have instead generated induced pluripotent stem cells (iPSCs) by directly delivering TFs into cells as fusion proteins with

© 2011 Elsevier Inc. All rights reserved.

*Correspondence to Department of Chemical Engineering, Stanford University, Stanford, California 94305; jswartz@stanford.edu; telephone: 650-723-5398; fax: 650-725-0555.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

protein transduction domains [6; 7; 8]. Though the somatic cells were successfully reprogrammed, the reprogramming efficiencies were low, approximately 0.001%.

The protein-based reprogramming approach suffers from low efficiencies, but the advantages of this approach suggest that recombinant TFs will be in demand. Researchers will seek to not only use and improve upon protein-based iPSC generation strategies but also to apply the same principles to control cell fate and directly convert somatic cells from one lineage to another. Given that there is growing interest in producing recombinant TFs for reprogramming applications, it is important to develop a rapid and quantitative way to evaluate TF product quantity, quality, and functionality *in vitro*.

One aspect of TF functionality that can be quickly measured *in vitro* is DNA binding. There are many widespread methods available for evaluating TF binding to its consensus sequence. Although the electrophoretic mobility shift assay (EMSA) is the most commonly used method for this purpose [9], it suffers from several significant drawbacks. It is tedious, time-consuming, difficult to adapt for higher throughput, requires large amounts of protein, and often requires radioactivity. Non-radioactive EMSA using fluorescently labeled DNA (Invitrogen) addresses the radiation issue, but the aforementioned drawbacks inherent to EMSA still apply. More importantly, the gel-based EMSA method does not directly provide quantitative estimates for apparent DNA binding affinity or capacity. For these reasons, we sought to develop a rapid and easy method for quantitatively evaluating the quality of recombinant TFs using low-cost, readily available equipment and components. Here, we present an assay method that directly provides quantitative evaluation of DNA-protein interactions without the need for explicit purification. The His₆-tagged recombinant TF is captured using Ni-NTA agarose. Then, the immobilized TF is incubated with a fluorescently-labeled DNA probe containing its respective consensus sequence. After washing away the unbound DNA, the level of fluorescence is measured. The fluorescence signal is directly proportional to the absolute amount of DNA bound to the TF. Using this information, we can determine the specificity of TF binding to its consensus sequence as well as measure the apparent affinity of the recombinant TF for its consensus sequence. Furthermore, if we use radiolabeled transcription factors, we can calculate a DNA to TF binding ratio to indicate the fraction of TFs that is “active,” or competent for DNA binding.

We validate the assay using two TFs that play important roles in regulating pluripotency, SOX2 and NANOG. We show that they specifically bind to their consensus DNA sequences and determine the apparent affinity of SOX2 and NANOG for their respective consensus sequences. By performing the assay with radiolabeled proteins, we can also measure the DNA to protein binding ratio. The inherent His₆ purification aspect of the protocol allows us to evaluate DNA binding activity from either purified or crude samples. Our assay format is highly flexible. Crude or purified samples can be used with or without radioactivity.

Materials and Methods

Expression vectors

The gene encoding for human SOX2 was amplified from pET24a-R9-SOX2 using PCR [10]. Flanking NdeI/NheI restriction sites and an N-terminal chloramphenicol acetyltransferase translation enhancer [11] followed by a His₆-tag were then added using PCR primers containing these features. Similarly, the gene encoding for human NANOG was amplified from pET24a-R9-NANOG using PCR [10], and NdeI/NheI restriction sites and a C-terminal His₆-tag were added. The PCR products were then digested with NdeI and NheI and ligated into pET24a (Novagen, Madison, WI) between the T7 promoter and terminator to form pET24a-CAT9-His₆-SOX2 and pET24a-NANOG-His₆. Thus, the SOX2 protein was expressed with an N-terminal His₆-tag while the NANOG was expressed with a

C-terminal His₆-tag. The ligation products were transformed into competent dH5α *E. coli* (Invitrogen, Carlsbad, CA). The resulting pET24a-based expression vectors from single colonies were verified by DNA sequencing. Milligram quantities of plasmid were isolated from *E. coli* cultures grown in Terrific Broth (Invitrogen) using Maxiprep and Gigaprep plasmid purification kits (QIAGEN, Valencia, CA) according to the manufacturer's instructions.

Cell-free protein synthesis

Cell-free protein synthesis (CFPS) was performed using the PANox-SP (PEP, Amino acids, Nicotinamide adenine dinucleotide (NAD), Oxalic acid, Spermidine, and Putrescine) cell-free system as previously described [12]. Minor changes to component concentrations are described below. All reagents were obtained from Sigma-Aldrich (St. Louis, MO) unless otherwise noted. The standard PANox-SP CFPS reaction mixture includes: 1.2 mM ATP, 0.85 mM each of GMP, UMP and CMP, 33 mM phosphoenol pyruvate (Roche Molecular Biochemicals, Indianapolis, IN), 175 mM potassium glutamate, 20 mM ammonium glutamate, 10 mM magnesium glutamate, 1.5 mM spermidine, 1.0 mM putrescine, 34 µg/mL folic acid, 171 µg/mL *E. coli* tRNA mixture (Roche Molecular Biochemicals, Indianapolis, IN), 13.3 µg/mL plasmid, 300 µg/mL T7 RNA polymerase (T7 RNAP) prepared as described [13], 2 mM of each of the 20 unlabeled amino acids, 0.33 mM NAD, 0.26 mM Coenzyme A (CoA), 2.7 mM sodium oxalate, 0.24 volumes of BL21Star(DE3) pGroESL *E. coli* S30 extract (see below for details), and with or without 6.25 µM l-[U-¹⁴C]-leucine (PerkinElmer, Waltham, MA). CFPS reactions were conducted in multiple 1-mL volumes in 6-well tissue culture plates (BD Falcon, Bedford, MA). Reactions were carried out at room temperature for 3 h. Following the CFPS reaction and a centrifugation step (20,800g for 15 minutes) to remove insoluble proteins, the supernatant containing soluble radioactive SOX2 or NANOG was used to perform the DNA binding assay. Production yields of SOX2 (37 kDa) and NANOG (36 kDa) were verified using liquid scintillation counting [14].

E. coli cell extract preparation

The *E. coli* S30 extract was prepared similarly to previously described extracts [10; 15]. Minor changes to the protocol are described below. BL21Star(DE3) (Invitrogen) harboring the pGroESL plasmid [16] was used to prepare the extract. Chemically competent BL21Star(DE3) was transformed with pGroESL and plated on LB agar with 34 µg/mL chloramphenicol. Single colonies were picked and 30% glycerol stocks of the BL21Star(DE3) pGroESL were made from cultures grown on defined media containing 34 µg/mL chloramphenicol. The glycerol stocks were then used to inoculate an 8 L high density fermentation culture using defined media [15] containing 34 µg/mL chloramphenicol. 0.5 mM IPTG was added at 6.0 OD₆₀₀ to induce both T7 RNAP expression and subsequent GroES/GroEL expression. The high density fermentation was harvested at 30 OD₆₀₀ and the cell extract, hereafter referred to as BL21Star(DE3) pGroESL, was prepared using procedures similar to those previously described [10]. Cells were washed 3× in S30 buffer and 0.8 mL of S30 buffer was added per gram of cell paste after the final wash. The cell suspension was then lysed by a single pass through an Emulsiflex C-50 (Avestin, Ottawa, ON, Canada) high pressure homogenizer at 17,500-25,000 psi. The homogenate was clarified by centrifugation at 30,000g at 4° C, 2× for 30 minutes each. The resulting pellets were discarded. The supernatant was incubated for 80 minutes at 37° C in the dark on a rotary shaker at 120 rpm and then clarified using a final centrifugation spin at 12,000g at 4° C for 20 minutes. After the final spin, the supernatant was flash-frozen and stored at -80° C.

Oligonucleotide substrates for the DNA binding assay

Single-stranded DNA oligonucleotides were purchased from Integrated DNA Technologies (IDT, Coralville, IA). Respective consensus sequences are underlined. The SOX2 cognate DNA probe contained the SOX2 consensus sequence identified as the HMG-binding site of the construct CR4 in POU5F1 and consisted of the sequence GCA GAG GAC AAA GGT GCC GTG [17]. The NANOG consensus sequence, TAATGG, was based on the Tcf3 enhancer. The NANOG cognate DNA probe sequence was ACC TGT TAA TGG GAG CGC [18]. All fluorescently-labeled DNA primers (forward and reverse) were modified at the 5' end with 6-carboxyfluorescein (6-FAM). The nonsense DNA probe sequence was GCC CGA TTA CTC TGT CCT. The forward and reverse complement oligonucleotide pairs were resuspended in 10 mM Tris-HCl, 50 mM NaCl, pH 8.0 at a concentration of 100 μ M. Oligonucleotides were annealed by heating for 5 minutes at 95° C and allowing the tubes to cool to room temperature on the benchtop.

DNA binding assay

The DNA binding assay is outlined in Figure 1. It involves three main steps: (1) protein capture, (2) DNA capture, and (3) DNA and protein quantification.

(1) Protein capture—100 μ L of a 50% Ni-NTA agarose bead suspension (QIAGEN) was added to each well of a 96-well filter plate (MSHVN4510, Millipore, Bedford, MA). The beads were washed 4 $\times$ with 200 μ L Load Buffer (LB, 300 mM NaCl, 10 mM imidazole, 50 mM sodium phosphate buffer pH 8.0) using a Millipore MultiScreen vacuum manifold assembly with pressures between 5-10 in Hg. Then, 10 μ L of 10 $\times$ concentrated LB was added to different volumes (3.3 to 90 μ L) of the soluble fraction of the crude CFPS reaction product containing His₆-tagged DNA-binding proteins of interest. H₂O was used to bring the volume up to 100 μ L. This 100 μ L load sample was added to each bead-containing well of the 96-well plate filter plate. The filter plate was covered and incubated for 2 h at 4° C on a titer plate shaker.

(2) DNA capture—After the 2 h incubation at 4° C, the beads were washed 2 $\times$ with 200 μ L of Wash Buffer (WB, 40 mM imidazole in LB) and 2 $\times$ with 200 μ L of Binding Buffer (BB, 20 mM HEPES-KOH pH 8.0, 50 mM KCl, 0.5mM DTT, 0.05mM EDTA, 1mM MgCl₂, 5% glycerol, 0.05% Tween20) using the vacuum manifold. Then, 100 μ L of different concentrations of fluorescently-labeled cognate DNA in BB was added to each well. The filter plate was then covered and incubated for 1 h at room temperature on a titer plate shaker.

(3) DNA and protein quantification—After the 1 h incubation at room temperature, the beads were washed 4 $\times$ with 200 μ L of PBS with 0.05% Tween20 (PBST) using the vacuum manifold. Then, 150 μ L of PBST was added to each well containing Ni-NTA beads. The filter plate was covered and incubated for 5 minutes at room temperature on a titer plate shaker. The Ni-NTA/PBST slurry was then transferred from the filter plate to a 96-well clear bottom black wall plate (Corning Costar, Corning, NY). Each well was excited at 485 nm and the fluorescence signal (RFU) was read at 535 nm using a Berthold Mithras LB 940 multimode plate reader (Berthold Technologies, Oak Ridge, TN). If the CFPS reactions had contained ¹⁴C-leucine, then the contents of each well of the 96-well plate were subsequently transferred to a scintillation vial and 5 mL of Beckman ReadySafe scintillation cocktail (Beckman, Fullerton, CA) was directly added to the bead/PBST slurry. Vials were vortexed for 10 seconds and radioactivity (cpm) was counted using a LS3801 liquid scintillation counter (Beckman).

Calibration measurements for fluorescently-labeled DNA and ^{14}C -leucine

Fluorescence and radiation calibration curves were constructed by measuring the RFU and cpm of known concentrations of fluorescently-labeled DNA and radioactive ^{14}C -leucine. The relationship between concentration and signal was determined by linear regression. These linear relationships were used to calculate the moles of DNA corresponding to a given RFU and moles of protein corresponding to a given cpm based upon the leucine content of the protein and the quantity of ^{14}C -leucine added to the 2mM concentration of non-radioactive leucine.

Competitive binding analysis

The amount of fluorescently-labeled cognate DNA and radioactive protein were both held constant. 10 μL of 10 $\times$ concentrated LB was added to 50 μL of the soluble fraction of the crude SOX2 or NANOG CFPS reaction product. The total volume was then brought up to 100 μL using 40 μL of H_2O . When incubated with 50 μL of the soluble SOX2 and NANOG final CFPS reaction mixtures (3.9 μM SOX2 and 5.9 μM NANOG; Supplementary Figure 1), 50 μL of Ni-NTA agarose beads captured approximately 8×10^{-11} moles of SOX2 and 4×10^{-11} moles NANOG (Supplementary Figure 2). After the products were captured by Ni-NTA, DNA binding was performed using a variety of competitor DNA probes. There were four main test cases: (1) fluorescent cognate, (2) cognate competitor, (3) nonsense competitor, and (4) no TF. All conditions contained 2 μM cognate DNA in a 100 μL volume. The cognate competitor and nonsense competitor conditions contained, in addition to the fluorescently-labeled cognate DNA probe, 10 μM of the non-labeled competitor cognate DNA probe and nonsense DNA probe, respectively. The no TF condition served as a negative control to show that the cognate DNA does not bind Ni-NTA agarose or any endogenous *E. coli* proteins that often co-purify with His₆-tagged proteins. For the no TF condition, a blank CFPS reaction mixture (no SOX2 or NANOG plasmid added to the reaction mixture) was incubated with Ni-NTA agarose and then assayed using 2 μM of fluorescently-labeled cognate DNA in a 100 μL volume.

Determination of apparent DNA binding affinity

The amount of fluorescently-labeled cognate DNA was varied while the amount of radioactive protein was held constant. 10 μL of 10 $\times$ concentrated LB was added to 50 μL of the soluble fraction of the crude SOX2 and NANOG CFPS reaction product. The total volume was then brought up to 100 μL using 40 μL of H_2O . Proteins were captured by the Ni-NTA agarose beads, and DNA binding was performed using serial dilutions of fluorescently-labeled cognate DNA (final concentrations of 2, 0.67, 0.22, 0.07, and 0.02 μM) in 100 μL volumes. 5-fold excess (relative to labeled cognate DNA) of non-labeled nonsense DNA was present in all scenarios. DNA concentrations were calculated using standard calibration curves described above. The quantity of bound cognate DNA (calculated from the fluorescence signal) was plotted against the concentration of free cognate DNA probe (calculated by measuring the fluorescence of the filtrate after the DNA incubation step). The binding affinity (K_D) of SOX2 and NANOG for its respective cognate DNA was calculated by fitting the data to a one-site saturation model accounting for non-

specific binding: $y = \frac{B_{\max}x}{K_D + x} + N_s x$, where y is the fluorescence signal, x is the concentration of unbound fluorescent DNA from the filtrate, N_s is the fitted non-specific binding constant, K_D is the apparent dissociation constant, and $B_{\max}$ is the number of available sites for DNA binding. This was performed using the Ligand Binding analysis package included as part of SigmaPlot11 (Systat Software, San Jose, CA).

Determination of DNA to protein binding ratio

The amount of fluorescently-labeled cognate DNA was held constant while the amount of radioactive protein was varied. 10 μ L of 10 $\times$ concentrated LB was added to 3.3, 10, 30, or 90 μ L of the soluble fraction of the crude radioactive SOX2 and NANOG CFPS reaction product for a total volume of 100 μ L. This corresponds to approximately 4.1×10^{-12} , 1.5×10^{-11} , 4.6×10^{-11} , 1.3×10^{-10} moles of SOX2 and 2×10^{-12} , 8×10^{-12} , 3×10^{-11} , 8×10^{-11} moles of NANOG captured by the Ni-NTA agarose beads (Supplementary Figure 2). After the proteins were captured by the Ni-NTA agarose beads, DNA binding was performed using 2 μ M fluorescently-labeled cognate DNA and 10 μ M non-fluorescently-labeled nonsense competitor DNA in BB in 100 μ L volumes. By incorporating 14 C-leucine into the protein, we can measure absolute amounts of both bound DNA and captured protein. Moles of cognate DNA, obtained from fluorescence signal, were plotted versus the moles of protein, obtained from radiation signal. Linear regression was performed to determine the DNA to protein binding ratio (slope).

Results

Protein production

We produced the His₆-tagged SOX2 and NANOG proteins using the PANOX-SP CFPS system. CFPS yielded approximately 250 μ g/mL of SOX2 and approximately 400 μ g/mL of NANOG, of which approximately half of each remained in the soluble fraction (Supplementary Figure 1).

SOX2 and NANOG bind specifically to their consensus DNA sequences

We characterized the specificity of the SOX2 and NANOG DNA binding interactions using competitive binding analysis. After capturing the His₆-tagged SOX2 and NANOG using Ni-NTA agarose beads, we incubated the immobilized SOX2 and NANOG proteins with their respective fluorescently-labeled cognate DNA probes with or without competitor probes. We observed a strong fluorescence signal when the immobilized TF probes were incubated with only their respective fluorescently-labeled cognate DNA (Figure 2). Co-incubation with excess non-labeled cognate DNA lowered the signal while co-incubation with excess non-labeled nonsense DNA did not significantly diminish the fluorescence signal. As a negative control, Ni-NTA agarose beads incubated with a blank CFPS reaction mixture that produced no TF (no plasmid added) were incubated with fluorescently-labeled cognate DNA. As expected, the fluorescence signal was very low, suggesting that DNA binding is specific.

SOX2 and NANOG apparent DNA binding affinity

This assay also serves as a convenient platform for quantifying a DNA-binding protein's apparent affinity for its cognate DNA. The 96-well plate format allows us to easily explore different binding conditions. The directly correlated RFU-to-DNA relationship also facilitates quantitative data analysis. Different concentrations of fluorescently-labeled cognate DNA (2, 0.67, 0.22, 0.07, and 0.02 μ M) were incubated with a fixed amount of SOX2 or NANOG. After accounting for ligand (DNA) depletion, the binding signal (moles of cognate DNA bound) was plotted against the concentration of free cognate DNA. The resulting curve was fitted to the standard model for one-site saturation with non-specific binding (Figure 3). The apparent affinity of SOX2 for its cognate DNA was determined to be 54.2 ± 9 nM while the apparent affinity of NANOG for its cognate DNA was determined to be 44.0 ± 6 nM.

Assay quantifies percentage of protein that is competent for DNA binding

This assay provides an additional layer of information if the protein is labeled with ^{14}C -leucine. In this work, we took advantage of the open nature of CFPS and radiolabeled the SOX2 and NANOG using ^{14}C -leucine incorporation. This allowed us to measure the absolute amount of protein adsorbed onto the beads. By labeling both DNA and protein, we can calculate a DNA to protein binding ratio. This ratio provides an estimate for the percentage of protein that can bind to its consensus sequence, or in other words, the percent “active.” To calculate the ratios for SOX2 and NANOG, we incubated different amounts of radiolabeled protein (4.1×10^{-12} , 1.5×10^{-11} , 4.6×10^{-11} , 1.3×10^{-10} moles of SOX2 and 2×10^{-12} , 8×10^{-12} , 3×10^{-11} , 8×10^{-11} moles of NANOG) with a fixed excess of fluorescently-labeled cognate DNA (2×10^{-10} moles). We then plotted the moles of adsorbed DNA against the moles of adsorbed protein and performed linear regression to determine the slope; i.e., the DNA to protein ratio (Figure 4). The data suggest that almost all of the SOX2 captured by the Ni-NTA bead binds its cognate DNA while approximately half of the NANOG captured by the Ni-NTA bead binds its cognate DNA. Alternatively, since NANOG is known to form a dimer [19], it is likely that each dimer binds only one DNA molecule. We confirmed that the maximal SOX2 ratio is approximately 1:1 and the maximal NANOG ratio is approximately 0.5:1 by varying the fluorescently-labeled cognate DNA concentration while holding the protein concentration constant (Figure 5). The matrix-immobilized protein binds additional cognate DNA until the number of available binding sites saturate. The point of saturation corresponds to the amount of protein that is competent for binding cognate DNA.

Discussion

One of the strengths of our assay is that it can supply much of the information provided by a standard EMSA without using radioactivity. Both the competitive binding analysis and determination of apparent DNA binding affinity can be performed without the use of radioactive protein. Traditional EMSA requires ^{32}P labeling of the DNA probe. In contrast, all of the key reagents required for this assay excluding the protein-of-interest are readily available. The Ni-NTA agarose and fluorescently-labeled oligonucleotides can be purchased and used without the need for end-user modification.

EMSA with subsequent densitometry has been used to estimate the binding affinities of SOX2 and NANOG for their respective consensus sequences. The K_D of SOX2 has been reported to be 15 nM [20] while the K_D of NANOG has been reported to be 26 nM [18]. The 54.2 ± 9 and 44.0 ± 6 nM values that we obtained are comparable to the literature values. The differences may be due to the use of different measurement modalities and binding buffer conditions (salt, temperature, etc.).

While the above-described information can be obtained without using radioactive materials, ^{14}C labeling takes advantage of our assay's full quantitative capabilities. By labeling the protein with radioactive amino acids, we can determine a DNA to protein binding ratio. The DNA to protein binding ratio provides insights into protein quality and suggests either that the NANOG may have protein folding problems or, more likely, that the dimer binds to a single DNA molecule. Many laboratories do not use radioactivity. In these instances, the protein concentration may be alternatively estimated using colorimetric assays. This can be performed on either purified protein or crude protein. If crude protein is used, densitometry can be performed on PAGE gel images to determine the percentage of the total protein that is the protein-of-interest.

The assay method we present is simple and versatile. It takes advantage of a common design feature for the preparation of recombinant proteins, the His₆-tag. The Ni-NTA capture step

is essentially a miniaturized purification, so the assay is capable of measuring the DNA binding parameters of either purified or crude protein samples. Performing the assay using a 96-well filter plate and vacuum manifold dramatically increases the speed and throughput. We obtained the repeated results less than 3 hours after incubating the crude protein samples with the Ni-NTA agarose beads. Alternatively, lower throughput versions of the assay can be performed using eppendorf tubes if vacuum manifolds are not available.

However, the assay does have its limitations. For instance, the need for the His₆-tag capture mechanism does not enable the assay to quantify the DNA binding activity of endogenous DNA-binding proteins. Similarly, if a protein is misfolded in such a way that the His₆-tag is not exposed for binding, then the assay will not yield binding information. However, the latter issue can be addressed. Firstly, both N-terminal and C-terminal extensions can be tested. If this is not effective, the affinity-based His₆-tagged capture and 96-well filter plate format is capable of screening protein folding conditions. For example, the assay could take advantage of the open environment of CFPS and be used to screen for CFPS production conditions that fold proteins in such a way that (1) the His₆-tag is exposed and (2) an optimal DNA to protein ratio is generated. Alternatively, the assay could be used to screen different refolding conditions from solubilized inclusion body starting material as many of the TFs of interest suffer from poor soluble expression [7; 10].

It is important to note that positive DNA binding activity does not automatically translate to transcriptional activity inside cells. TFs are complex biomolecules that interact not only with their consensus sequences, but also with other TFs, polymerases, and epigenetic enzymes. By showing that a TF binds its consensus DNA sequence, we show that the DNA binding domain is exposed and folded correctly; nothing is indicated about the folding state of the rest of the protein. However, if a TF does not bind its consensus DNA, then there is a good likelihood that it will not be biologically active. Therefore, DNA binding can serve as a prerequisite for deciding whether to proceed with more time- and resource-consuming biological assays.

In summary, the protocol presented in this work is widely applicable for evaluating the DNA binding of a variety of recombinant DNA-binding proteins in a rapid and quantitative format. By capturing the recombinant His₆-tagged TF using Ni-NTA agarose and incubating it with fluorescently-labeled cognate DNA, we can directly correlate the fluorescence signal with the amount of DNA bound. We validated this assay using two recombinant TFs, SOX2 and NANOG. We showed that they bind specifically to their consensus sequences and determined the apparent binding affinity of SOX2 and NANOG for their consensus sequences. Our assay is highly flexible, giving the end-user the option to use pure or crude samples with or without radioactivity. Though radioactivity is not required, labeling the protein products does provide an additional layer of information, the DNA to TF binding ratio. We used this assay to calculate this ratio for SOX2 and NANOG in order to estimate the fraction of recombinant TFs that is competent for DNA binding and/or the binding stoichiometry.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Dong-Myung Kim, Jieun Lee, Bertrand Lui, and Cem Albayrak for helpful discussions and Dong-Myung Kim for his careful reading of the manuscript and helpful comments. This work was supported by the National Institutes of Health [5U01HL100397]. WCY is a recipient of a National Defense Science and Engineering Graduate Fellowship and National Science Foundation Graduate Fellowship.

References

- [1]. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007; 131:861–72. [PubMed: 18035408]
- [2]. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin, Thomson JA. Induced pluripotent stem cell lines derived from human somatic cells. *Science*. 2007; 318:1917–20. [PubMed: 18029452]
- [3]. Stehling-Sun S, Dade J, Nutt SL, DeKoter RP, Camargo FD. Regulation of lymphoid versus myeloid fate ‘choice’ by the transcription factor Mef2c. *Nat Immunol*. 2009; 10:289–96. [PubMed: 19169261]
- [4]. Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Sudhof TC, Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*. 463:1035–41. [PubMed: 20107439]
- [5]. Maherali N, Hochedlinger K. Guidelines and techniques for the generation of induced pluripotent stem cells. *Cell Stem Cell*. 2008; 3:595–605. [PubMed: 19041776]
- [6]. Kim D, Kim CH, Moon JI, Chung YG, Chang MY, Han BS, Ko S, Yang E, Cha KY, Lanza R, Kim KS. Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell*. 2009; 4:472–6. [PubMed: 19481515]
- [7]. Zhou H, Wu S, Joo JY, Zhu S, Han DW, Lin T, Trauger S, Bien G, Yao S, Zhu Y, Siuzdak G, Scholer HR, Duan L, Ding S. Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell*. 2009; 4:381–4. [PubMed: 19398399]
- [8]. Cho HJ, Lee CS, Kwon YW, Paek JS, Lee SH, Hur J, Lee EJ, Roh TY, Chu IS, Leem SH, Kim Y, Kang HJ, Park YB, Kim HS. Induction of pluripotent stem cells from adult somatic cells by protein-based reprogramming without genetic manipulation. *Blood*.
- [9]. Hellman LM, Fried MG. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc*. 2007; 2:1849–61. [PubMed: 17703195]
- [10]. Yang WC, Patel KG, Lee J, Ghebremariam YT, Wong HE, Cooke JP, Swartz JR. Cell-free production of transducible transcription factors for nuclear reprogramming. *Biotechnol Bioeng*. 2009; 104:1047–58. [PubMed: 19718703]
- [11]. Son JM, Ahn JH, Hwang MY, Park CG, Choi CY, Kim DM. Enhancing the efficiency of cell-free protein synthesis through the polymerase-chain-reaction-based addition of a translation enhancer sequence and the in situ removal of the extra amino acid residues. *Anal Biochem*. 2006; 351:187–92. [PubMed: 16430851]
- [12]. Jewett MC, Swartz JR. Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol Bioeng*. 2004; 86:19–26. [PubMed: 15007837]
- [13]. Li Y, Wang E, Wang Y. A modified procedure for fast purification of T7 RNA polymerase. *Protein Expr Purif*. 1999; 16:355–8. [PubMed: 10419832]
- [14]. Calhoun KA, Swartz JR. An economical method for cell-free protein synthesis using glucose and nucleoside monophosphates. *Biotechnol Prog*. 2005; 21:1146–53. [PubMed: 16080695]
- [15]. Zawada J, Swartz J. Maintaining rapid growth in moderate-density Escherichia coli fermentations. *Biotechnol Bioeng*. 2005; 89:407–15. [PubMed: 15635610]
- [16]. Goloubinoff P, Gatenby AA, Lorimer GH. GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose biphosphate carboxylase oligomers in Escherichia coli. *Nature*. 1989; 337:44–7. [PubMed: 2562907]
- [17]. Chew JL, Loh YH, Zhang W, Chen X, Tam WL, Yeap LS, Li P, Ang YS, Lim B, Robson P, Ng HH. Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol*. 2005; 25:6031–46. [PubMed: 15988017]
- [18]. Jauch R, Ng CK, Saikatendu KS, Stevens RC, Kolatkar PR. Crystal structure and DNA binding of the homeodomain of the stem cell transcription factor Nanog. *J Mol Biol*. 2008; 376:758–70. [PubMed: 18177668]
- [19]. Wang J, Levasseur DN, Orkin SH. Requirement of Nanog dimerization for stem cell self-renewal and pluripotency. *Proc Natl Acad Sci U S A*. 2008; 105:6326–31. [PubMed: 18436640]

- [20]. Scaffidi P, Bianchi ME. Spatially precise DNA bending is an essential activity of the sox2 transcription factor. *J Biol Chem*. 2001; 276:47296–302. [PubMed: 11584012]

Figure 1. Schematic of the DNA binding assay

DNA binding is quantified by fluorescence and protein concentration is quantified by radioactivity. Alternatively, radioactivity can be omitted to perform a non-radioactive version of this DNA binding assay.

Figure 2. SOX2 and NANOG bind specifically to their cognate DNA sequences

Competitive binding analyses were performed using fluorescently-labeled cognate DNA probes, non-fluorescent cognate DNA probes, and nonsense competitive DNA probes. SOX2 cognate DNA probe: GCA GAG GAC AAA GGT GCC GTG; NANOG cognate DNA probe: ACC TGT TAA TGG GAG CGC; Nonsense competitor DNA probe: GCC CGA TTA CTC TGT CCT. Consensus sequences are underlined. When incubated with fluorescently-labeled DNA without the presence of competitors, DNA-protein binding was observed. Co-incubation of fluorescently-labeled cognate DNA with excess non-labeled cognate competitor probe reduced the DNA binding signal. Co-incubation of fluorescently-labeled cognate DNA with excess non-labeled nonsense competitor probe resulted in the return of the DNA binding signal. A no protein negative control resulted in a very low DNA binding signal, indicating DNA binding was dependent on the presence of the DNA-binding protein of interest. The amounts of fluorescently-labeled cognate probe (2 μ M) and protein in each sample were kept constant (8×10^{-11} mol SOX2, 4×10^{-11} mol NANOG). Error bars indicate standard deviation, $n = 3$.

Figure 3. Determination of SOX2 (A) and NANOG (B) cognate DNA apparent binding affinity Different amounts of fluorescently-labeled cognate DNA were incubated with fixed amounts of radiolabeled SOX2 (8×10^{-11} mol) and NANOG (4×10^{-11} mol). A binding isotherm was constructed by accounting for cognate DNA (ligand) depletion. The apparent binding affinity was obtained using a Systat Sigmaplot11 non-linear regression package that fit the data to a one-site saturation model accounting for non-specific binding interaction. The estimated binding affinities are indicated. Error bars indicate standard deviation, $n = 3$.

Figure 4. DNA to protein binding ratio

Different amounts of radiolabeled protein were incubated with a fixed amount of fluorescently-labeled cognate DNA (2 μ M). After the DNA incubation, the bead-associated fluorescence was determined using fluorometry and the bead-associated radioactivity was counted using scintillation counting. The DNA to protein binding ratio was obtained from the slope m of the linear regression $y = mx + b$. Results suggest a 1:1 ratio and 0.5:1 ratio of DNA to protein for SOX2 and NANOG, respectively. Error bars indicate standard deviation, $n = 3$.

Figure 5. Matrix-immobilized protein binds free cognate DNA until binding site saturation

Different amounts of fluorescently-labeled cognate DNA were incubated with fixed amounts of radiolabeled SOX2 (8×10^{-11} mol) and NANOG (4×10^{-11} mol). The amount of fluorescently-labeled DNA bound to beads at the end of the assay was plotted against the amount of fluorescently-labeled DNA made available for binding at the start of the assay. SOX2 binding of its cognate DNA increases linearly until it plateaus at the number of SOX2 molecules available for binding, supporting the 1:1 ratio from Figure 4. NANOG binding of its cognate DNA increases linearly until it reaches approximately 50% of the number of NANOG molecules available for binding, supporting the 0.5:1 ratio from Figure 4. Together, these data suggest that approximately 100% of the CFPS SOX2 is competent for DNA binding while either approximately 50% of the CFPS NANOG is competent for DNA binding or two molecules of NANOG bind one molecule of DNA.