Published in final edited form as:

Methods Mol Biol. 2014 ; 1123: 127–134. doi:10.1007/978-1-62703-968-0_10.

Rapid determination of homing endonuclease DNA binding specificity profile

Lei Zhao¹ and Barry L. Stoddard^{2,*}

¹Residency program in Pathology, University of Chicago, 5841 S. Maryland Ave, Chicago, IL, 60637

²Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N. Seattle WA 98109

Summary

Evaluating the binding specificity and identifying the most preferred target sequence for a homing endonuclease often represents a key step during its characterization, engineering and application for genome engineering. This chapter describes a high throughput, fluorescence based, competition binding assay which can be used to measure the relative binding affinities of the homing endonuclease to a large number of DNA target site variants in a single experiment. The base preference at each position of the target sequence can be quantitated based on this assay, and the overall binding specificity of the enzyme can thereby be determined and compared with its cleavage specificity.

Keywords

Homing endonuclease; DNA binding protein; Specificity; High throughput; Fluorescence-based; Competition binding assay; Base preference

1. Introduction

The assay described in this chapter can be used to determine the binding specificity profile of a homing endonuclease, by simultaneously measuring the binding affinities of the homing endonuclease to a large number of DNA sequences, each of which varies only slightly from the wild type target sequence. This assay was developed based on previously described methods which use a fluorescence-based, microplate format to facilitate high throughput analyses that include large numbers of replicate measurements for each DNA target ^{1–2}. The main improvement of the assay described in this chapter is the employment of competition binding strategy, which not only minimizes the cost required for fluorescent labeling of the DNA target, but also greatly improves sensitivity and reproducibility ³.

An overview of the assay is illustrated in Figure 1. Briefly, the assay is performed in 96 well microplates; an independent binding assay for one of the sequence variants (referred to below as a “test”) is conducted in each individual well. The recombinant homing endonuclease with an N- or C-terminal poly-histidine tag is first immobilized onto nickel coated microplates. The immobilized homing endonuclease is then incubated with a small amount of fluorescence-labeled wild-type DNA duplex, which is presented along with an excess of unlabeled competitor DNA. A different unlabeled DNA sequence variant is added

*To whom correspondence should be addressed: bstoddard@fhcrc.org phone: 1-206-667-4031 fax: 1-206-667-3331, Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N. A3-025, Seattle WA 98109.

alongside the labeled wild-type DNA target in each well. The relative binding affinities of the homing endonuclease to each individual sequence variant are measured by determining the ability of each DNA sequence variant to compete for binding against the fluorescently labeled wild-type target sequence (see note ¹). After incubation of the DNA mixture with the homing endonuclease, the unbound DNA (labeled or unlabeled) is washed off and the amount of remaining labeled wild type DNA can be quantified by a fluorescence reader.

The fluorescent signal in each test well is inversely proportional to the binding affinity of the homing endonuclease to the corresponding sequence variant. In order to compute the relative binding affinity, two additional types of tests must be included in the experiment. The first additional test should have wild type DNA sequence as added unlabeled competitor, and the second additional test should include a random unlabeled DNA sequence (a completely randomized sequence used as a negative control to account for competition by a nonspecific DNA sequence) which is also used to measure binding competition by a nonspecific DNA sequence. The relative binding affinity of each sequence variant to that of the wild-type target can be calculated based on the readings of above three types of tests. The base preference at each position of target DNA of the homing endonuclease can then be derived from the relative binding affinities of a comprehensive matrix of sequence variants.

2. Material

1. Ni-NTA HisSorb Plates, white
2. Poly dI-dC: dissolve in TBS buffer at 1 microgram per microliter, aliquot, warm to 45° C for 5 min, cool, and freeze.
3. Synthesized unlabeled sequence variants (see note ²)
4. Synthesized unlabeled wild type sequence
5. Synthesized random DNA sequence
6. 5'Cy3TM labeled wild type sequence
7. 6×-His-tagged recombinant homing endonuclease
8. 5×TBS: 250 mM Tris-HCl, pH 7.5, 750 mM NaCl
9. TBS/BSA: 1× TBS with 0.2% BSA
10. TBS/Tween-20: 1× TBS with 0.05% Tween-20
11. Binding buffer: 1× TBS with 0.02 mg/ml poly dI-dC, 10mM CaCl₂
12. PCR H₂O: milli-Q H₂O filtered by 0.22 micron pore size sterile filters

¹The fluorescently labeled sequence used in this assay is often the wild type target sequence for the homing endonuclease. However, any other sequence with high binding affinity (sub nano-molar range) to the homing endonuclease can be used as the standard sequence. As long as the absolute affinity of the standard sequence is known, the binding affinities of other sequence variants can be inferred following above instructions.

²We usually add a short flanking random sequence (such as “AAAAAA”) to both ends of the DNA substrates to ensure effective binding of the homing endonuclease to its target DNA. This is based on the observation that some homing endonucleases make non-specific contacts to the phosphate backbone of the flanking sequences from structural studies.

3. Methods

3.1 Preparation of DNA duplexes

1. Dissolve the synthesized fluorescence-labeled oligos into PCR H₂O at 200 millimolar concentration. Mix equal amount of top strand and bottom strand, which leads to a final concentration of 100 millimolar for the annealed duplex. Use a PCR machine to anneal the oligos (Heat to 90° C for 10 minutes; slowly cool to RT. Slow cooling is important to avoid hairpin formation; an overnight annealing reaction in a PCR machine is recommended).
2. The synthesized sequence variants can be dissolved in solution at a normalized concentration by the manufacturer upon request. Mix equal amount of top strand and bottom strand in PCR tubes. Anneal the duplex as described in step 1.

3.2 Immobilization of the homing endonuclease

1. Dilute the his-tagged homing endonuclease sample to 100 nM protein concentration with TBS/BSA buffer, and then load 200 µl protein into each microwell of the HisSorb plates.
2. Incubate at room temperature or in a cold room (depending on the stability of the homing endonuclease under investigation) for at least 2 hours. For more efficient binding, incubate on a plate shaker.
3. After incubation, discard the solutions in the plates by quickly inverting the plate and taping on paper towels gently to dry.
4. Wash the plates four times with TBS/Tween-20. Allow soaking for 1 min between washes.
5. Discard the buffer after final wash.

3.3 Binding

Each micro-well contains an independent binding assay, in which an unlabeled competitor DNA competes for binding against the fluorescently labeled wild type DNA. The unlabeled competitor may be one of the following three types: (1) "WT" (the unlabeled wild-type DNA site competing against itself); (2) "Random" (a completely randomized sequence used as a negative control to account for competition by a nonspecific DNA sequence); or (3) one of the DNA target site variants. We use a sequence matrix (i,j) to identify individual sequence variant, in which i stands for the position in the target DNA, and j represents the base substitution at that position. All three types of unlabeled DNA competitor are treated independently in all steps.

1. Mix labeled wild type DNA with unlabeled competitor DNA in binding buffer. The final concentration of the labeled DNA should be 100 nanomolar and that of the unlabeled competitor is 3 micromolar (see note ³).
2. Apply 200 microliters of the mixture to each well and incubate for at least 2 hours.
3. Wash the plates with TBS for four times. Dry the plate on paper towel as described above.
4. Apply 200 microliters TBS to each well after final wash.

³To minimize mixing and transferring, a 10 × oligonucleotide mix (with 1 µM labeled DNA and 30 µM unlabeled DNA) can be made first. Load 20 microliters of this 10× oligo mix into each microwell with immobilized protein and then dilute with 180 microliters of binding buffer.

3.4 Fluorescence reading

Measure the retained fluorescence in each well using a SpectraMax® M5/M5^e microplate reader (excitation: 510nm, emission: 565nm, cutoff: 550).

3.5 Data analysis

The measurements of the retained fluorescent signal (F) for each mismatch sequence variant [$F(i,j)$] are then converted to relative binding affinities as compared to the wild-type target site using the relationship $rK_a(i,j) = [(F(\text{random}) - F(i,j)) \times F(\text{wt})] / [F(i,j) \times (F(\text{random}) - F(\text{wt}))]$. This formula gives a close approximation of the binding affinity of each sequence variant (Figure 2).

The relative base-pair preference of the homing endonuclease under investigation at each position can then be calculated using the relationship: $\text{BP}(i,j) = K_a(i,j) / \sum_{j=A,C,G,T} K_a(i,j)$. An example of the full calculation of $K_a(i,j)$ and $\text{BP}(i,j)$ from raw fluorescence intensity measurements is provided in the note ⁴.

3.6 Remarks about this method and data

It is agreed that for most if not all homing endonucleases, that a considerable fraction of the overall cleavage specificity of the enzyme is derived from the actual catalytic action of the enzyme (i.e. during the course of DNA cutting) rather than during initial recognition and binding of the target site. In other words, the cleavage specificity of most homing endonucleases is substantially higher than is their binding affinity. However, the determination of a homing endonuclease's binding specificity can be quite important, for at least two reasons. First, the mere binding of a homing endonuclease at a large number of sites with significant affinity can have notable effects on genomic fidelity and cell viability, particularly for dividing cell populations. Second, it has become clear that for at least some homing endonucleases, that the basis of cleavage specificity across the target site can be

⁴Example data conversion (see Figure 2)

$$rK_a(i,j) = [(F(\text{random}) - F(i,j)) \times F(\text{wt})] / [F(i,j) \times (F(\text{random}) - F(\text{wt}))],$$

$$\text{BP}(i,j) = K_a(i,j) / \sum_{j=A,C,G,T} K_a(i,j)$$

(for i =position at the target DNA, j =A,C,G,T)

are used for converting raw data to relative binding affinity (rK_a) and base preference (BP) as described in methods. Below are details of data conversion at one example basepair substitution at position -5:

a. Raw fluorescence signals:

$$F(-5, A)=313.20, F(-5, C)=329.75, F(-5, T)=270.00, F(-5, G)=F(\text{wt})=84.17, F(\text{random})=342.66$$

b. Conversion to relative K_a (rK_a) values:

$$rK_a(-5, A) = [(F(\text{random}) - F(-5, A)) \times F(\text{wt})] / [F(-5, A) \times (F(\text{random}) - F(\text{wt}))] = [(342.66 - 313.20) \times 84.17] / [313.20 \times (342.66 - 84.17)] = 0.03$$

$$\text{Using the same equation, } rK_a(-5, C) = 0.01, rK_a(-5, T) = 0.08, rK_a(-5, G) = rK_a(\text{wt}) = 1$$

c. Conversion to Base Preference (BP):

$$\text{BP}(-5, A) = K_a(-5, A) / \sum_{j=A,C,G,T} K_a(-5, j) = K_a(-5, A) / (K_a(-5, A) + K_a(-5, C) + K_a(-5, T) + K_a(-5, G)) = 0.03 / (0.03 + 0.01 + 0.08 + 1) = 2.7\%$$

$$\text{BP}(-5, C) = K_a(-5, C) / \sum_{j=A,C,G,T} K_a(-5, j) = K_a(-5, C) / (K_a(-5, A) + K_a(-5, C) + K_a(-5, T) + K_a(-5, G)) = 0.01 / (0.03 + 0.01 + 0.08 + 1) = 0.9\%$$

$$\text{BP}(-5, T) = K_a(-5, T) / \sum_{j=A,C,G,T} K_a(-5, j) = K_a(-5, T) / (K_a(-5, A) + K_a(-5, C) + K_a(-5, T) + K_a(-5, G)) = 0.08 / (0.03 + 0.01 + 0.08 + 1) = 7.2\%$$

$$\text{BP}(-5, G) = K_a(-5, G) / \sum_{j=A,C,G,T} K_a(-5, j) = K_a(-5, G) / (K_a(-5, A) + K_a(-5, C) + K_a(-5, T) + K_a(-5, G)) = 1 / (0.03 + 0.01 + 0.08 + 1) = 89.3\%$$

unevenly distributed between binding (K_M) and catalysis (k_{cat})⁴. In either case, a quantitative understanding of binding specificity and affinity can provide a solid foundation for the manipulation, engineering and application of homing endonucleases for targeted gene modification and genome engineering.

References

1. Zhang ZR, Palfrey D, Nagel DA, Lambert PA, Jessop RA, Santos AF, Hine AV. Fluorescent microplate-based analysis of protein-DNA interactions. I: Immobilized protein. *BioTechniques*. 2003; 35:980–982. 984, 986. [PubMed: 14628672]
2. Hallikas O, Taipale J. High-throughput assay for determining specificity and affinity of protein-DNA binding interactions. *Nature protocols*. 2006; 1:215–222.
3. Zhao L, Pellenz S, Stoddard BL. Activity and specificity of the bacterial PD-(D/E)XK homing endonuclease I-Ssp6803I. *J Mol Biol*. 2009; 385(5):1498–1510. [PubMed: 19038269]
4. Thyme SB, Jarjour J, Takeuchi R, Havranek JJ, Ashworth J, Scharenberg AM, Stoddard BL, Baker D. Exploitation of binding energy for catalysis and design. *Nature*. 2009; 461:1300–1304. [PubMed: 19865174]

Figure 1.

High-throughput fluorescence competition-binding assay.

Histidine-tagged protein is immobilized onto nickel-coated 96-well microplates. A mixture of fluorescently labeled DNA duplex containing the wild-type target site and one of the unlabeled sequence variants is incubated with the protein, allowing determination of the ability of each sequence to compete for binding against the wild-type DNA target. In order to calculate relative affinities, the experiment is also conducted using unlabeled wild-type DNA target and a nonspecific, random DNA sequence.

a **b** **Figure 2.**

Panel a: Example data conversion. The raw fluorescent reading, relative Ka (rKa) and Base Preference (BP) are plotted in the left, middle and right panel, respectively. Refer to the text for detail. **Panel b:** Binding specificity profile for wild-type I-AniI (a LAGLIDADG homing endonuclease). The effect of a single basepair substitution at each position across a 20 basepair sequence corresponding to the endonuclease's wild-type target is shown. The binding affinity of the enzyme to its wild-type target is normalized and denoted as "1"; the relative affinity (rKa) of the same protein to each of 60 separate DNA targets (each of which contains a single basepair substitution at one position, from basepair -10 to basepair +1) is displayed. The height of the bars denotes the effect of each possible substitution at each position, as described in the text and illustrated in the middle panel of **2a** above. The error bars represent standard deviations for 3 independent measurements. Note that the enzyme displays considerable binding specificity at positions -10 to -3, and then considerably lower binding specificity at positions +3 to +10. In contrast, the enzyme displays considerable cleavage specificity across the entire target site. See reference 4 for more details ⁴.