Electrophoretic Mobility Shift Assays with GFP-Tagged Proteins (GFP-EMSA)

Running title: GFP-based assay for investigation of nucleic acid-binding proteins

Sorenson AE[1]

Schaeffer PM[1]

[1] Molecular and Cell Biology, College of Public Health, Medical and Veterinary Sciences, James Cook University, Douglas QLD 4811, Australia

*Corresponding author

Associate Professor Patrick Schaeffer

Phone: +61 (0)7 4781 4448

Email: patrick.schaeffer@jcu.edu.au

Address: The Science Place, Building 142, University Drive, James Cook University, Douglas, Queensland 4811, Australia

i. Summary

The electrophoretic mobility shift assay (EMSA) is commonly used for the study of nucleic acid binding proteins. The technique can be used to demonstrate that a protein is binding to RNA or DNA through visualization of a shift in electrophoretic mobility of the nucleic acid band. A major disadvantage of the EMSA is that it does not always provide an absolute certitude that the band shift is due to the protein under scrutiny, as contaminants in the sample could also cause the band shift. Here we describe a variation of the standard EMSA allowing to visualize with added certitude, the co-localized band shifts of a GFP-tagged protein binding to its cognate nucleic acid target sequence stained with an intercalator such as GelRed. Herein, we present an illustrative protocol of this useful technique called GFP-EMSA along with specific notes on its advantages and limitations.

ii. Key Words

Green fluorescent protein, nucleic acid binding, DNA binding, RNA binding, electrophoretic mobility shift assay, band shift assay, gel shift assay.

1 Introduction

The interaction of proteins with nucleic acids is fundamental to transcription, post-transcriptional modification, translation, DNA replication, repair and recombination. Due to their central function in such a wide variety of essential cellular processes

and their role in disease development, nucleic acid-binding proteins have become desired therapeutic targets [1,2]. Characterization of essential protein:nucleic acid interactions is important for both fundamental research purposes, and as a basis for drug development. The EMSA has long been an important staple tool in the detection and characterisation of nucleic acid binding proteins and their target sequences [3]. In a typical EMSA, target nucleic acid sequences can be labeled with a fluorescent probe, radioactive isotopes nuclei, or stained for detection. The presence of the cognate binding protein is indirectly visualized through its effect on the mobility of its target sequence [3]. GFP-EMSA is a unique technique that allows the concomitant detection of the protein and its target sequence. GFP-EMSA can be performed with relatively safe nucleic acid binding dyes post-electrophoresis. This technique is particularly useful for the definitive characterization of protein:nucleic acid interactions [4-8].

2 Materials

2.1 Reagents

1. GFP-tagged nucleic acid binding protein (see **Note 1**).

2. Oligonucleotides (see **Note 2**).

3. Appropriate fluorescent label/intercalating dye (see **Note 3**).

4. Annealing buffer: 20 mM Tris-HCl pH 8.0, 100 mM NaCl.

5. Binding buffer (dependent on cofactors and strength of interaction if known, see **Note 4**; if unknown: 10 mM Tris pH 8.0, 1% glycerol, 50 mM NaCl).

6. 50% glycerol.

7. Loading buffer (any 5-6x loading buffer with dye to indicate progression).

2.2 Gel Electrophoresis

1. Molecular biology grade agarose (see **Note 5**).

2. Glass flask or bottle sufficient for volume of gel required (dependent on number of reactions desired and electrophoresis equipment available).

3. Microwave.

4. Gel electrophoresis systems, gel molds, combs, electrophoresis tank and power pack (see **Note 6**).

5. Running buffer: Tris-borate-EDTA (TBE), 89 mM Tris, 89 mM boric acid, 2 mM EDTA [9].

2.3 Gel Imaging

1. UV transilluminator and camera with adequate filter or Gel documentation system capable of emitting UV, blue and green light, with appropriate filters for red/green imaging (e.g. Syngene G:BOX).

2. Nucleic acid stain if required for oligonucleotide detection (e.g. Biotium GelRed).

3 Methods

3.1 Reagent preparation

The following sample protocol is typically applied for a GFP-tagged protein binding to a double stranded (ds) DNA target sequence with post-electrophoretic GelRed staining.

1. Combine equal volumes of both oligonucleotides (1 μM each) in annealing buffer and heat for 2 min at 80°C.

2. Allow to cool slowly to room temperature (RT). The volume to prepare will depend on the number of reactions required. Reactions are typically 10 μL volume containing 1 μM GFP-tagged DNA binding protein and target DNA sequence at desired ratio.

3.2 Gel electrophoresis

1. Prepare a thin 1% agarose gel in TBE without nucleic acid stain (see **Note 7**).

2. Mix equal volumes (5 μl each) of GFP-tagged protein (1 μM) with ds DNA (1 μM) in binding buffer (see **Note 8**).

3. Include two control reactions that contain protein only and ds DNA only for comparison of their respective electrophoretic mobility (see **Note 9**).

4. Incubate reactions at RT for 10 min.

5. Add 2.5 μL 50% glycerol to reactions if binding buffer contained less than 5% glycerol (see **Note 10**).

6. Load 10 μL samples into the TBE gel with an additional lane of generic coloured loading buffer to enable visualization of progression (see **Note 11**).

7. Subject gel to electrophoresis at 80 V for sufficient time to separate bands (see **Note 12**).

3.3 Gel imaging

1. The gel can be exposed to UV or blue light to excite the GFP chromophore (**see Note 13**). If a GFP fusion protein band shift is apparent then incubate gel with 1x GelRed in water for 30 min with rocking at RT (see **Note 14**).

2. Expose gel with adequate excitation wavelengths (i.e. for GFP and GelRed) and capture emitted light with a CCD camera using appropriate filters (see **Note 15**, Fig. 1).

3. Analyse image results (see **Note 16**)

[Fig 1 near here]

4 Notes

1. The GFP-tagged protein can be produced recombinantly using typical *E. coli* expression systems [14,15]. Alternatively, the GFP-tagged protein can be purified from prokaryotic or eukaryotic systems where it is being used as a reporter of gene expression [16] and localisation *in vivo* [17]. GFP-tagged DNA binding proteins have been expressed *in vivo* in all domains of life [18-21] . Production using *in vitro* transcription-translation is also an option [22]. We have successfully applied the technique with cycle 3 GFP [10,11] as well as mCherry [12] fusion proteins and

expect it to work with proteins tagged with any fluorescent protein as long as appropriate excitation wavelengths and filters are used and care has been taken to choose an appropriate DNA fluorescent dye (i.e. the fluorescent protein and DNA dye do not have overlapping emission spectra or excitation maxima). We typically use proteins with a C-terminal GFP tag, however if this is likely to interfere with binding, an N-terminal tag could be used instead. Dual fluorophore EMSAs can be performed if the imaging system has multiple excitation wavelengths and filters.

2. GFP-EMSA can be performed with DNA and RNA in single or ds form. Many oligonucleotide binding proteins will require flanking sequences to enable optimal protein binding. We typically use 5 nt flanking sequences both at 5′ and 3′ ends of the target sequence. Nucleic acid binding proteins have varying affinity and specificity for nucleic acids. The length of oligonucleotides will be dependent on the binding protein. We have successfully performed GFP-EMSA with oligonucleotides ranging from 20-115 bp. It is worth noting that the longer the nucleic acid the less shift in mobility will be observed with a stoichiometric binding mode.

3. Generic fluorescent intercalating dyes can be used such as GelRed (Biotium) or ethidium bromide. Other more specific nucleic acid binding dyes, such as SYBR Green I or II can be used if required. Alternatively an oligonucleotide can be labelled with a fluorescent dye (e.g. Cy3 or Cy5).

4. GFP-EMSA has been performed with a wide range of buffers and salt concentrations. Thus, binding buffer composition will be dependent on the DNA binding conditions that are to be tested (e.g. salt or pH dependence [13]). A typical

binding reaction can be performed in PBS, but other buffer systems such as Tris can be used, and salt concentrations can also be altered depending on application. Some proteins may require the addition of divalent cations such as Mg^{2+} or other cofactors such as ATP.

5. High quality agarose has very low background fluorescence and is sufficiently transparent for performing GFP-EMSA. However, polyacrylamide (8-10% gel) can be used as an alternative when higher sensitivity is required.

6. It is important to use gel electrophoresis equipment that has not been in contact with intercalating nucleic acid dyes, or it will need to be cleaned thoroughly after previous use to remove any trace of such dyes. This is because intercalating nucleic acid dyes interfere with protein binding.

7. A thin gel improves resolution and reduces background. Pour a gel as thin as possible with sufficient depth to accommodate a 10 μL reaction volume. Polyacrylamide gels can be used (8-12%) in either horizontal or vertical formats. With vertical systems, the unbound protein might not be visible depending on its pI.

8. The amount of protein required can be dependent on the quantum yield of the fluorophore and sensitivity of the imaging equipment. We find that 1 μM GFP-tagged protein is generally sufficient for fluorescence capture. If the stoichiometry of the complex is not known, it can be estimated by trialing different ratios of proteins and oligonucleotides.

9. Control reactions will demonstrate the mobility of each component in isolation and enable visualization of any band shifts produced by the interaction between the protein and nucleic acid (Fig. 1).

10. Glycerol is added to increase the density of the sample to assist with gel loading. It also increases protein stability [23].

11. A coloured loading dye is added to a separate well to use as a marker during electrophoresis. It is not added to the samples themselves as it may migrate at the same position as the oligonucleotide/s and/or protein, making imaging difficult.

12. The time taken for electrophoresis will be dependent on the size of the oligonucleotides as well as percentage of agarose and buffer system. For most proteins 30-60 min will be sufficient. Longer oligonucleotides may require longer electrophoresis for sufficient separation. GFP-EMSA can be paused at any time to check the fluorescent protein band progression using an adequate excitation wavelength. For cycle 3 GFP simply excite with UV or blue light using a 525 nm filter. A standard "black light" flashlight is often sufficient to check progression during electrophoresis. If using a polyacrylamide gel for electrophoresis, use chilled TBE, keep cool during electrophoresis and do not exceed 8 V/cm.

13. Cycle 3 GFP has bimodal absorption spectra and can be excited with light ranging from UV to blue (350-480 nm) with an emission maximum at 509 nm.

14. GelRed does not require de-staining. The gel can be rinsed with water to reduce background if necessary. However, this has not been an issue with GFP-EMSA using

the protocol described. The same procedure can be used for other dyes such as SYBR Green II (use at 1:10,000 in TBE).

15. This protocol is based on the Syngene G:BOX XRQ. Other imaging systems can be used if the light sources and filters are similar. For imaging cycle 3 GFP alone expose gel to either blue Epi LED (465 nm) or long wave UV light (365 nm) using a FILT525 filter (516-539 nm) (Fig.1). For imaging GelRed stained nucleic acid, expose gel to medium wave UV (MW UV) light (302 nm) using a UV06 filter (572-625 nm) or FILT605M filter (594-610 nm)(Fig. 1). For imaging mCherry, expose gel to green Epi LED (525 nm) using a FILT605M filter (594-610 nm). Possible fluorophores will be dependent on the available equipment.

16. Integrate fluorescent bands with image analysis software (e.g. ImageJ). Overlay images to illustrate and confirm co-localization of GFP and DNA bands in the gel.

References

1. Miyagi T, Shiotani B, Miyoshi R, Yamamoto T, Oka T, Umezawa K, Ochiya T, Takano M, Tahara H (2014) DSE-FRET: A new anticancer drug screening assay for DNA binding proteins. Cancer Sci 105 (7):870-874. doi:10.1111/cas.12420

2. Alonso N, Guillen R, Chambers JW, Leng F (2015) A rapid and sensitive high-throughput screening method to identify compounds targeting protein-nucleic acids interactions. Nucleic Acids Res 43 (8):e52. doi:10.1093/nar/gkv069

3. Hellman LM, Fried MG (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. Nat Protoc 2 (8):1849-1861. doi:10.1038/nprot.2007.249

4. Moreau MJ, Morin I, Schaeffer PM (2010) Quantitative determination of protein stability and ligand binding using a green fluorescent protein reporter system. Mol Biosyst 6 (7):1285-1292. doi:10.1039/c002001j

5. Askin SP, Bond TEH, Schaeffer PM (2016) Green fluorescent protein-based assays for high-throughput functional characterization and ligand-binding studies of biotin protein ligase. Anal Methods-Uk 8 (2):418-424

6. Dahdah DB, Morin I, Moreau MJ, Dixon NE, Schaeffer PM (2009) Site-specific covalent attachment of DNA to proteins using a photoactivatable Tus-Ter complex. Chem Commun (Camb) (21):3050-3052. doi:10.1039/b900905a

7. Johnston EB, Kamath SD, Lopata AL, Schaeffer PM (2014) Tus-Ter-lock immuno-PCR assays for the sensitive detection of tropomyosin-specific IgE antibodies. Bioanalysis 6 (4):465-476. doi:10.4155/bio.13.315

8. Morin I, Schaeffer PM (2012) Combining RNA-DNA swapping and quantitative polymerase chain reaction for the detection of influenza A nucleoprotein. Anal Biochem 420 (2):121-126. doi:10.1016/j.ab.2011.09.009

9. Green MR, Sambrook J (2012) Molecular Cloning: A Laboratory Manual, vol 3. 4th edn. Cold Springs Hoarbour Laboratory Press,

10. Fukuda H, Arai M, Kuwajima K (2000) Folding of green fluorescent protein and the cycle3 mutant. Biochemistry 39 (39):12025-12032

11. Wang L, Xie J, Deniz AA, Schultz PG (2003) Unnatural amino acid mutagenesis of green fluorescent protein. J Org Chem 68 (1):174-176. doi:10.1021/jo026570u

12. Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, Tsien RY (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat Biotechnol 22 (12):1567-1572. doi:10.1038/nbt1037

13. Moreau MJ, Schaeffer PM (2013) Dissecting the salt dependence of the Tus-Ter protein-DNA complexes by high-throughput differential scanning fluorimetry of a GFP-tagged Tus. Mol Biosyst 9 (12):3146-3154. doi:10.1039/c3mb70426b

14. Tabor S, Richardson CC (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc Natl Acad Sci U S A 82 (4):1074-1078

15. Dubendorff JW, Studier FW (1991) Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. J Mol Biol 219 (1):45-59

16. Chalfie M (1995) Green fluorescent protein. Photochem Photobiol 62 (4):651-656

17. Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA (2010) Fluorescent proteins and their applications in imaging living cells and tissues. Physiol Rev 90 (3):1103-1163. doi:10.1152/physrev.00038.2009

18. Straight AF, Belmont AS, Robinett CC, Murray AW (1996) GFP tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. Curr Biol 6 (12):1599-1608

19. Mortusewicz O, Schermelleh L, Walter J, Cardoso MC, Leonhardt H (2005) Recruitment of DNA methyltransferase I to DNA repair sites. Proc Natl Acad Sci U S A 102 (25):8905-8909. doi:10.1073/pnas.0501034102

20. Lindhout BI, Fransz P, Tessadori F, Meckel T, Hooykaas PJ, van der Zaal BJ (2007) Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins. Nucleic Acids Res 35 (16):e107. doi:10.1093/nar/gkm618

21. Delpech F, Collien Y, Mahou P, Beaurepaire E, Myllykallio H, Lestini R (2018) Snapshots of archaeal DNA replication and repair in living cells using super-resolution imaging. Nucleic Acids Res. doi:10.1093/nar/gky829

22. Spirin AS, Swartz JR (2008) Cell-free Protein Synthesis: Methods and Protocols. John Wiley & Sons,

23. Vagenende V, Yap MG, Trout BL (2009) Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. Biochemistry 48 (46):11084-11096. doi:10.1021/bi900649t

24. Beckett D, Kovaleva E, Schatz PJ (1999) A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. Protein Sci 8 (4):921-929. doi:10.1110/ps.8.4.921

25. Eisenstein E, Beckett D (1999) Dimerization of the Escherichia coli biotin repressor: corepressor function in protein assembly. Biochemistry 38 (40):13077-13084

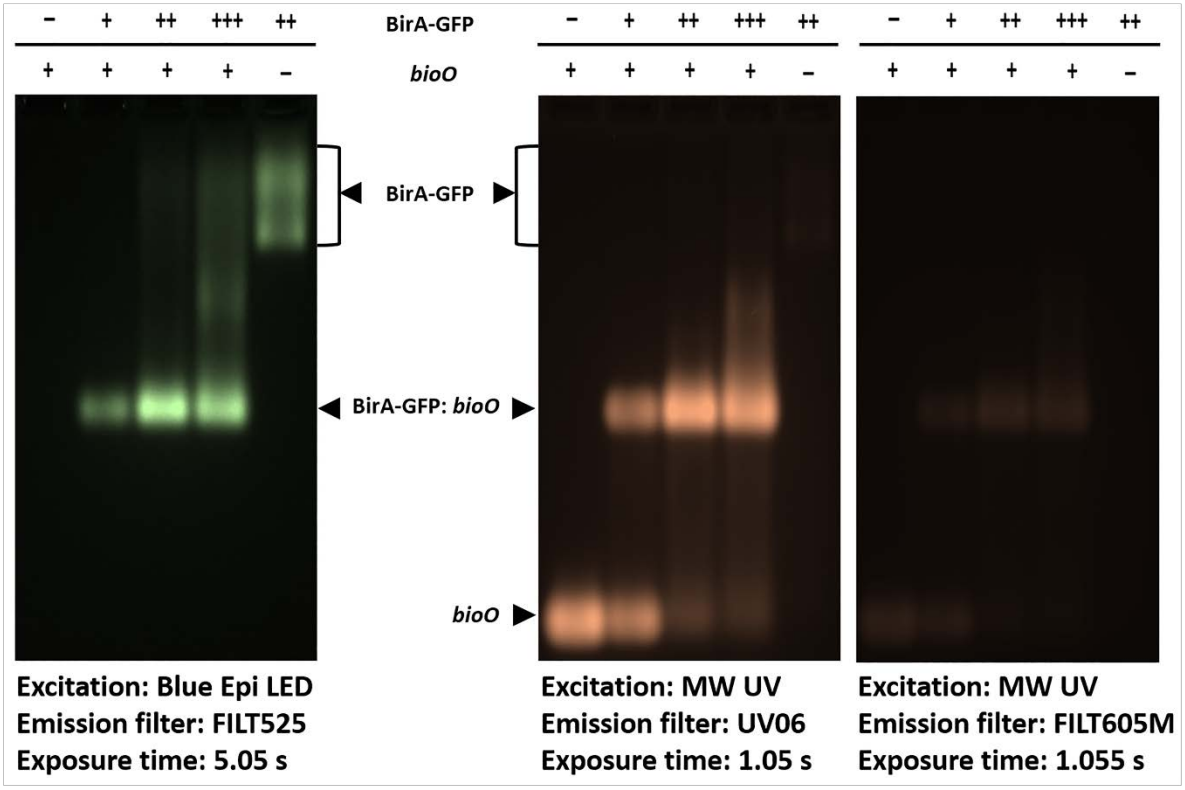

Figure 1: Example GFP-EMSA

In *E. coli*, the bifunctional BirA protein is responsible for biotinylation of proteins and regulation of biotin synthesis through binding to the biotin operator (*bioO*) [24,25]. In this example 5 µL *bioO* (0.8 µM) was combined with different concentrations of BirA-GFP (0.4, 0.8 and 1.2 µM) in PBS, pH 7.4 with 0.6 mM ATP and biotin and 3 mM $MgCl_2$ to demonstrate complex formation, similarly as described previously [5]. A band shift can be seen with BirA-GFP in the presence of *bioO* (+). Images were captured using a G:BOX Chemi XRQ (excitation, filter and exposure times detailed below each image) and recoloured to highlight differential visualization of GFP (green) and GelRed stained DNA (orange). MW UV denotes medium range UV (302 nm).