Quantitative approaches to monitor protein–nucleic acid interactions using fluorescent probes

JOHN M. PAGANO, CARINA C. CLINGMAN, and SEAN P. RYDER
Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA

ABSTRACT
Sequence-specific recognition of nucleic acids by proteins is required for nearly every aspect of gene expression. Quantitative binding experiments are a useful tool to measure the ability of a protein to distinguish between multiple sequences. Here, we describe the use of fluorophore-labeled oligonucleotide probes to quantitatively monitor protein/nucleic acid interactions. We review two complementary experimental methods, fluorescence polarization and fluorescence electrophoretic mobility shift assays, that enable the quantitative measurement of binding affinity. We also present two strategies for post-synthetic end-labeling of DNA or RNA oligonucleotides with fluorescent dyes. The approaches discussed here are efficient and sensitive, providing a safe and accessible alternative to the more commonly used radio-isotopic methods.

Keywords: fluorescence polarization; EMSA; gel shift; fluorophore; post-synthetic labeling; affinity; protein–RNA interaction

INTRODUCTION
Myriad cellular processes require sequence specific recognition of a nucleic acid by a protein. Transcription factors bind to specific DNA elements to enhance or repress transcription at a defined locus (Deplancke et al. 2006; Barrasa et al. 2007; Carrera and Treisman 2008; Noyes et al. 2008). RNA-binding proteins coordinate translation, mRNA localization and stability, and pre-mRNA splicing through association with defined sequences in target transcripts (Varnum et al. 1991; Johnstone and Lasko 2001; Nilsen 2002; Jurica and Moore 2003; Sing and Varcarcel 2005; Iwasaki et al. 2009). As such, it is critically important to understand the basic mechanisms by which DNA- and RNA-binding proteins identify their appropriate target sequences.

There are many ways a protein can recognize a DNA or RNA sequence. Most DNA-binding proteins rely on variations in the pattern of hydrogen bond acceptors and donors in the major groove to achieve sequence-specific recognition (Seeman et al. 1976). Others identify specific sequences through their relative flexibility that induces distortions in the DNA duplex (McClarin et al. 1986; Kim et al. 1993). RNA-binding proteins can recognize sequences through readout of the hydrogen bonding patterns of the bases on the Watson-Crick and Hoogsteen faces, or through shape-specific recognition of RNA secondary and tertiary structure (Howe et al. 1994; Välegard et al. 1994; Hudson et al. 2004; Hall 2005).

To identify the determinants of specificity, it is often necessary to measure the quantitative parameters of binding, such as relative binding affinity and stoichiometry (Schaup et al. 1970; Samuels et al. 1994). There are a number of methods that can be used to monitor protein/nucleic acid complexes (Wong and Lohman 1993; Royer and Scarlata 2008; Ryder et al. 2008). One of the more prominent techniques is the electrophoretic mobility shift assay (EMSA) (Dahlberg et al. 1969; Ryder et al. 2008). EMSA is a powerful technique that allows the visualization of the interaction between a protein and a labeled DNA or RNA probe. Equilibration reactions that include a fixed concentration of a 32P-radiolabeled oligonucleotide probe and varying amounts of protein are run on a native polyacrylamide gel to separate bound from free nucleic acid. Gels are dried and then exposed to film or a phosphorimager plate in order to separate the bound probe as a function of protein concentration. These data can be fit to determine the apparent equilibrium dissociation constant.

The use of radioactive material has the advantage of making the assay extremely sensitive. Very small amounts of 32P can be detected using common equipment. Another advantage is that the labeled and unlabeled probes are chemically similar, reducing the chance that the binding reaction is perturbed.
due to indirect effects caused by the label. However, radiolabeled probes also have several disadvantages, including safety, environmental, and regulatory challenges. Moreover, radioactive labels are costly, and due to the short half-life of $^{32}$P, probes have to be frequently relabeled and radioisotope stocks reordered.

Due to the challenges associated with using radioisotopes, fluorescent probes have become a favorable alternative in many biochemical assays (Royer and Scarlata 2008; Le'Filly and Royer 1993; Aviv et al. 2003; Ryder and Williamson 2004; Hardin and Batey 2006; Besse et al. 2009; Kohn et al. 2010). With modern instrumentation, the detection of fluorescent probes now rivals that of $^{32}$P-labeled probes. In this review, we present two methods to generate end-labeled fluorescent probes using commercially available fluorescent dyes and simple aqueous phase chemical reactions that can be performed in any laboratory (Reines and Cantor 1974; Czworkowski et al. 1991). In addition, we present two complementary approaches to measure the affinity of a protein for labeled DNA and RNA probes that rely on different physical properties of the complex—fluorescence polarization (FP) and fluorescence EMSA (F-EMSA) (Hellman and Fried 2007; Pagano et al. 2007; Ryder et al. 2008). Both assays can be applied to the same equilibration reactions because the FP measurements do not destroy the sample. Furthermore, the F-EMSA method enables analysis of the gels directly following electrophoresis, avoiding the need to dry and expose the phosphorimagery screens, thereby reducing the time required to perform the assay.

We have used the approach described herein to study four RNA-binding proteins required for germline development and/or embryogenesis in the nematode Caenorhabditis elegans (Pagano et al. 2007, 2009; Farley et al. 2008) and two proteins that regulate oligodendrocyte differentiation in vertebrates (SP Ryder, unpubl.). Others have used it to study the association of bacteriophage P7 coat protein and vertebrate ZBP variants to their cognate RNA sequences (Chao et al. 2008, 2010). These results show that the method is applicable to a wide variety of protein–RNA complexes.

**LABELING STRATEGIES**

**5’-End labeling of DNA and RNA oligonucleotides**

Chemically synthesized single-stranded DNA and RNA oligonucleotides can be efficiently post-synthetically labeled at the 5’-end using a two-step semi-enzymatic synthesis strategy described by Czworkowski and coworkers (Fig. 1A; Czworkowski et al. 1991). In the first step, T4 polynucleotide kinase (PNK) is used to transfer the $\gamma$-phosphorothioate group from 5’-OH to the phosphorothioate group on the probe in HEPES (pH 7.4)-buffered water to create a covalent adduct between the oligonucleotide and the fluorophore. Unreacted dye is removed by ethanol precipitation followed by a size exclusion spin column. Analytical denaturing polyacrylamide gel electrophoresis is used to assess purity. The reaction yield is measured, comparing the absorbance at 491 nm (fluorescein) to that at 260 nm (nucleotides and fluorescein) (Fig. 1B). We typically observe 80%–95% labeling efficiency. The reaction is simple and efficient and does not require pre-existing chemical modification of the DNA or RNA oligonucleotide. The reactions can be performed at the bench using standard laboratory plasticware. Fume hoods and specialized glass are not required. DNA probes labeled in this way can be used as primers in primer extension, PCR, or RT-PCR assays; the label does not interfere with subsequent enzymatic applications. The reaction is compatible with other iodoacetamide fluorescent dyes, including 7-diethylamino-3-((4’-(iodoacetyl)amino)phenyl)-4-methylcoumarin (DCIA) and tetramethylrhodamine-5-iodoacetamide, enabling a wide variety of downstream applications in several colors. Moreover, the approach is not limited to chemically synthesized DNA or RNA oligonucleotides. In principal, any substrate of PNK can be labeled in this manner.

**3’-End labeling of RNA**

In 1974, Reines and Cantor described a simple synthetic method to conjugate fluorescent dyes to RNA molecules at the 3’-end (Reines and Cantor 1974). This approach takes advantage of the vicinal hydroxyls present only at the

---

**FIGURE 1.** 5’-End labeling of DNA and RNA oligonucleotides. (A) The 5’-end labeling strategy involves reacting unphosphorylated DNA or RNA with T4 PNK and ATP$\gamma$S to yield a thiol reactive group on the 5’-end. 5-Iodoacetamidofluorescein (5-IAF) is then reacted with the thiol group. (B) An example of an ultraviolet spectrum of a DNA oligonucleotide that was labeled with 5-IAF on the 5’-end. The DNA absorbs at 260 nM, and the fluorescein label absorbs at 492 nM.
3’-end of ribonucleotides to achieve specific labeling without the need to protect other sites in the RNA. In the first step of the reaction, the RNA is incubated with sodium periodate (NaIO₄) in a mildly acidic sodium acetate buffer (pH 5.1) in order to oxidize the vicinal hydroxyls to aldehydes (Fig. 2A). Following an ethanol precipitation step, the oxidized RNA is then reacted with fluorescein-5-thiosemicarbazide (FTSC) in order to generate a covalent thiosemicarbazone linkage with the RNA. The product can then be reduced with sodium cyanoborohydride to limit chemical reversibility, but this is not usually necessary. The labeled RNA is resolved from unreacted dye by ethanol precipitation followed by a size exclusion spin column. The products are analyzed as described above in the 5’-end labeling procedure (Fig. 2B).

As above, the reaction can be performed at the bench and does not require prior modification of the RNA oligonucleotide. Unless sodium cyanoborohydride is used, fume hoods and special glassware are not required. Any molecule with vicinal hydroxyls can be labeled this way, including ribonucleotides and longer RNA molecules produced by in vitro transcription. This is an important caveat; if in vitro transcribed RNA is used as the substrate, care must be taken to purify the RNA away from the other nucleotides present in the transcription reaction. Several aldehyde reactive dyes are available. The chemistry is the same for any thiosemicarbazide or hydrazide, including the commercially available Alexa dyes (Invitrogen) in a wide variety of colors, and the useful affinity tag biotinamidocaproyl hydrazide (BACH). As with the prior strategy, labeling efficiency is usually >80%.

There are, of course, other suitable labeling strategies, including body labeling by PCR or in vitro transcription with fluorescent nucleotide analogs (Besse et al. 2009; Kohn et al. 2010). Incorporation of amino, azido, or thiol groups during chemical synthesis of DNA or RNA oligonucleotides that can react with a wide variety of fluorescent dyes may also be used. These strategies are efficient and can be used to make labeled material for the quantitative applications outlined below. However, we prefer the end-labeling methods as any oligonucleotide can be labeled without incorporating reactive moieties during synthesis, and certain applications require an end label rather than random incorporation of a body label. Thus, the methods presented here remove a barrier to the use of fluorescent dyes by making post-synthetic labeling as convenient as the current strategies used to radiolabel nucleic acids.

QUANTITATIVE FLUORESCENCE METHODS TO MONITOR PROTEIN–NUCLEIC ACID INTERACTIONS

FP and F-EMSA can be used to determine the affinity and specificity of a DNA- or RNA-binding protein (Fig. 3). Both assays rely on different physical properties of the complex. Each require recombinant protein that has been purified to >95% homogeneity and a fluorescently labeled DNA or RNA probe. Both assays conveniently use the same experimental setup; thus it is possible to perform both assays using the same series of binding reactions. Since the FP is nondestructive, F-EMSA can be performed immediately following the polarization measurement.

FP assays

FP is a useful technique to study the thermodynamic and kinetic properties of a protein–nucleic acid interaction (LeTilly and Royer 1993; Aviv et al. 2003; Ryder and Williamson 2004). FP takes advantage of the change in the tumbling properties of a fluorescent ligand upon binding to a larger macromolecule: in our case, a fluorescent DNA or RNA probe and a nucleic acid–binding protein. Polarized light is used to excite the fluorophore, and emitted light is measured in planes both parallel and perpendicular to the plane of excitation. A labeled probe that is free in solution tumbles more rapidly because of its comparatively low molecular weight, leading to the depolarization of the emitted light. However, tumbling is reduced when the nucleic acid is bound to protein, causing an increase in polarization. The maximal extent of the increase is dependent on the size of the protein, the size of the labeled probe, and the lifetime of the fluorophore.

To measure the apparent affinity of the protein–nucleic acid complex, a series of equilibration reactions are set up with varying concentrations of protein and fixed trace amounts of labeled probe. After equilibration, polarization...
is determined using a fluorometer or a fluorescence plate reader equipped with polarizers. The extent of polarization is influenced by the relative concentration of each fluorescent species in the equilibration reaction. As such, polarization is directly proportional to the fraction of bound probe at each protein concentration. It is also important to note that if higher-order species are formed or if there is unreacted dye in the equilibration reaction, these will contribute to the apparent polarization value as well. The effective polarization, expressed in units of millipolarization (mP), is related to the fluorescence intensity (I) in the parallel (para) and perpendicular (perp) planes by Equation 1. Most plate readers capable of FP measurements report this value automatically.

\[ mP = 1000 \times \frac{[\text{I} \text{para} - \text{I} \text{perp}]}{[\text{I} \text{para} + \text{I} \text{perp}]} \]  

The data are then plotted as a function of protein concentration and fit to the Hill Equation 2 in order to determine the apparent equilibrium dissociation constant (K_d) and the apparent cooperativity (n) between a macromolecule (P_i) and its ligand (Fig. 4). The equation presented here uses the maximum (m) and base (b) signals as normalization factors that represent the polarization values at the upper and lower asymptote of the titration (Hill 1910).

\[ \phi = b + (m - b)\left(\frac{1}{1 + (K_d/[P_i])^n}\right)^2 \]  

Other parameters that can be measured by this approach include the stoichiometry of the complex, determined by repeating the equilibration reactions at elevated probe concentration and fitting mP as a function of molar ratio to a quadratic equation derived by Rambo and Doudna (2004). FP is especially amenable to kinetic measurements, as the change in polarization can be measured in real time by many instruments.

The primary limitation of FP is that it requires a relatively small labeled ligand, usually ≤10 kDa when using fluorescein dyes. The lifetime of the fluorophore defines the size limit of the ligand. If the labeled probe cannot efficiently depolarize the emitted light within the lifetime of the fluorophore, it is not possible to measure the change in depolarization caused by binding to a protein. Other factors that contribute to the size limit include the shape and flexibility of the ligand. We typically use this approach for single-stranded RNA probes that are ≤30 nucleotides (nt). Smaller probes may be necessary for double-stranded DNA or RNA. The use of aliphatic linkers between the nucleic acid and the fluorophore should be avoided to minimize the “propeller effect,” which decouples fluorophore rotation from the rotational correlation time of the labeled DNA or RNA (Lynch et al. 1997; Wu et al. 1997). It is also important to control for fluorescence quenching. This is easily determined by monitoring fluorescence intensity in addition to anisotropy as a function of protein concentration. If quenching occurs, it may be necessary to try the label at the other end of the nucleic acid or to extend the nucleic acid beyond the footprint of the associating protein.

**Electrophoretic mobility shift assays**

F-EMSA is a complementary approach to measure the affinity of an interaction between a nucleic acid and a protein.
This method relies on the ability of a protein to influence the migration of a labeled nucleic acid through a native polyacrylamide or agarose gel (Hellman and Fried 2007; Ryder et al. 2008). The migration depends upon the length, and as a result the overall charge, as well as the shape of the nucleic acid. If association of a protein perturbs these parameters, it can change the rate of migration, enabling separation of bound and free probe. As with FP, a series of equilibration reactions are set up such that protein concentration varies while a trace amount of labeled probe remains at the same concentration in each reaction. After equilibration, the reactions are loaded onto the gel and subjected to electrophoresis in order to separate bound from free RNA. Then, the fraction of bound RNA is determined and fit as a function of protein concentration to the Hill equation as described above. The primary advantage of this approach is that it enables visualization of both free and bound probe. Thus, if multiple proteins bind to the DNA or RNA sequence, this usually manifests as multiple shifted species. Moreover, if the probe has degraded, this can be assessed as well. The disadvantage of this approach is that it does not provide a true equilibrium measurement. Complexes with fast off rates will dissociate in the time it takes to load the sample, perturbing the apparent affinity.

Traditionally, EMSA is performed with radioactively labeled probes in vertical gel equipment (Hellman and Fried 2007; Ryder et al. 2008). We have observed that fluorescent probes provide sufficient sensitivity for use in experiments that require low nanomolar concentrations of the labeled oligo. To achieve sufficient sensitivity, we adapted the EMSA approach to horizontal submarine polyacrylamide gels, which enables loading of more sample per well compared to standard vertical gel equipment. Moreover, the horizontal format enables multiplexing of experiments as more than one comb can be used to create wells in the gel matrix. Immediately following electrophoresis, the gel is imaged using a fluorescent-capable phosphorimagery equipped with the appropriate laser to excite the fluorophore (Fuji FLA-5000 or similar) (Fig. 5). This approach does not require drying the gel or exposing the gel to a phosphorimagery screen, which saves time relative to experiments performed with radioactive probes.

The data obtained by F-EMSA produce comparable results to radioactive experiments. For example, we determined the affinity of recombinant QKI RNA-binding domain for its high affinity binding site in the Mbp 3’-UTR by F-EMSA (Supplemental Fig. 1). The apparent equilibrium dissociation constant is 26 ± 14 nM, identical to the value previously determined by radioactive EMSA (25 ± 4 nM) (Ryder and Williamson 2004). However, it is worth noting that the apparent cooperativity seems greater by F-EMSA than by radioactive EMSA (n_{app} = 1.8 ± 0.4 versus 1.0 ± 0.2), which could be due to the increase in RNA concentration needed to detect binding, or to limitations of detection at the edge of the binding transition. Caution is warranted when interpreting the Hill coefficient by F-EMSA; we suggest comparing the Hill obtained by F-EMSA to values obtained by another approach, such as FP, to increase confidence in this parameter. In this case, the K_{d, app} determined by FP is 22 ± 9 nM, and the n_{app} is 1.1 ± 0.1.

The primary disadvantage of F-EMSA is that the minimum amount of substrate needed to make an observation depends upon fluorophore concentration, labeling efficiency, quantum yield, and instrument sensitivity. In our experience, 2–4 nM fluorescent end-labeled nucleic acid is needed to achieve sufficient signal-to-noise for a quantitative measurement. Thus, for proteins that bind to their target sequences with an apparent K_{d} <20–40 nM, the trace assumption associated with fitting the Hill equation will be violated. In such cases, it may be necessary to fit the data to a quadratic equation (Equation 3, where R is labeled nucleic acid concentration, P is the total protein concentration, and other symbols are labeled as in Equation 2) or use a substrate with more labeled nucleotides, such as an in vitro transcript that has been body-labeled with fluorescent nucleotide analog. Failure to do so will lead to over-estimation of the equilibrium dissociation constant and Hill coefficient.

\[
\phi = b + (m - b) \times \frac{R + P + K_d - \sqrt{(R + P + K_d)^2 - (4RP)}}{2R}
\]

**Concluding remarks**

This review presents a fluorescent alternative to common radiographic EMSA. The approach enables quantitative...
assessment of protein–nucleic acid binding affinity, specificity, and stoichiometry. Moreover, it improves upon standard radiographic methods in that two independent measurements—each of which depends on a different physical property of the protein–nucleic acid complex—can be obtained using the same reactions. The method uses standard equipment and instrumentation that is widely available in research settings. It provides a viable alternative when radioactive measurements are complicated by restrictive regulation. The approach should be generally useful to anyone interested in quantitative analysis of protein–nucleic acid interactions.

SUPPLEMENTAL MATERIAL

Supplemental material can be found at http://www.rnajournal.org.

ACKNOWLEDGMENTS

We thank Dr. Ruth Zearfoss and Brian Farley for helpful comments concerning this manuscript, and William Flaherty for technical assistance. This work was supported by a National Institutes of Health Grant GM081422 to S.P.R.

REFERENCES


Pagano et al.


Quantitative approaches to monitor protein–nucleic acid interactions using fluorescent probes

John M. Pagano, Carina C. Clingman and Sean P. Ryder

RNA 2011 17: 14-20 originally published online November 22, 2010
Access the most recent version at doi:10.1261/rna.2428111

Supplemental Material
http://rnajournal.cshlp.org/content/suppl/2010/11/16/rna.2428111.DC1

References
This article cites 41 articles, 12 of which can be accessed free at:
http://rnajournal.cshlp.org/content/17/1/14.full.html#ref-list-1

License

Email Alerts Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

To subscribe to RNA go to:
http://rnajournal.cshlp.org/subscriptions