# Probing Small-Molecule Microarrays with Tagged Proteins in Cell Lysates

Marius S. Pop,[1,2] Dina Wassaf,[1] and Angela N. Koehler[1,3]

[1]Broad Institute of Harvard and MIT, Cambridge, Massachusetts
[2]The Dana-Farber Cancer Institute, Boston, Massachusetts
[3]Koch Institute for Integrative Cancer Research at MIT, Cambridge, Massachusetts

The technique of small-molecule microarray (SMM) screening is based on the ability of small molecules to bind to various soluble proteins. This type of interaction is easily detected by the presence of a fluorescence signal produced by labeled antibodies that specifically recognize a unique sequence (tag) present on the target protein. The fluorescent signal intensity values are determined based on signal-to-noise ratios (SNRs). SMM screening is a high-throughput, unbiased method that can rapidly identify novel direct ligands for various protein targets. This binding-based assay format is generally applicable to most proteins, but it is especially useful for protein targets that do not possess an enzymatic activity. SMMs enable screening a protein in a purified form or in the context of a cellular lysate, likely providing a more physiologically relevant screening environment. © 2014 by John Wiley & Sons, Inc.

Keywords: small-molecule microarray • binding assay • cellular lysate • fluorescence • protein complex

**How to cite this article:**
*Curr. Protoc. Chem. Biol.* 6:209-220.
doi: 10.1002/9780470559277.ch140101

## INTRODUCTION

This article describes the general procedure for small-molecule microarray (SMM) screening whereby the primary goal is to identify putative binders to a given protein target or protein complex. Several probe discoveries enabled by SMMs have been recently reviewed (Hong et al., 2014). The process for manufacturing microarrays also has been reviewed previously (Uttamchandani et al., 2004; Walsh and Chang, 2004; Shi et al., 2010; Uttamchandani and Yao, 2010; Vegas and Koehler, 2010) and is not the focus of this protocol. The reader is referred to previous step-by-step protocols focused on the production of SMMs (Bradner et al., 2006; Casalena et al., 2012).

The SMM assay is based on the ability of compounds corresponding to various types of chemical structures to directly bind to soluble proteins, with specificity that depends on the properties of the small molecule as well as the three-dimensional structure of the protein surface. Highly specific antibodies that recognize a unique tag sequence engineered at either terminus of the target protein identify putative interactions between compounds and targets. The detection is based on a fluorescent signal generated by dyes (e.g., Cy3, Cy5) linked to the primary antibody. Alternatively, a labeled secondary antibody can be used to bind unlabeled primary antibodies. The fluorescent signal is recorded with a microarray scanner that generates an image file, and the signal intensities are analyzed using various methods, but the most common analysis methods focus on signal-to-noise ratios.



*Current Protocols in Chemical Biology* 6:209-220, December 2014
Published online December 2014 in Wiley Online Library (wileyonlinelibrary.com).
doi: 10.1002/9780470559277.ch140101

The protocol presented herein (see Basic Protocol) describes a screening approach involving a tagged target protein in the context of a cell lysate, and has been successfully applied previously to targets recalcitrant to ligand discovery such as transcriptional regulators and extracellular factors (Stanton et al., 2009; Miyazaki et al., 2010; Pop et al., 2014). Lysate-based binding assays enable researchers to discover different types of probes from screens using purified proteins. In lysates, target proteins can be closer to their native folded states and often have post-translational modifications that might be important for their activity. The lysate format can provide the appropriate environment for target proteins to remain engaged with other protein or nucleic acid binding partners, which in turn enhances their stability. Additionally, screening using lysates enables researchers to identify multiple types of assay positives. Some compounds may bind the target of interest directly, while others may bind to a binding-partner protein, which may have an impact on the function of the target of interest. Obtaining a pool of assay positives that might interact with various domains of a target protein or to binding partners is very powerful when dealing with novel targets, particularly those that function in biomolecular complexes. Screening using lysates may also enable methods to identify novel modes of modulating conventional targets. A key technical advantage of lysate screens over purified protein screens relates to soluble proteins present in the lysate that serve as blocking agents, dramatically reducing background signals and leading to data of overall higher quality. While the protocol described herein is focused on recombinantly tagged and expressed protein, the lysate approach can also involve detection of small molecules that bind to endogenous proteins, when expression levels are sufficient and direct antibodies are available.

## STRATEGIC PLANNING

This protocol is compatible with SMMs manufactured using multiple surface capture strategies or formats and assumes that printed SMMs are available at the time of the screening. The screener should consider whether the surface chemistry used for SMM production results in a linkage that is stable to cellular esterases prior to executing the screen. Refer to the previously published protocols on SMM production for lysate-compatible arrays (Casalena et al., 2012) or reviews that thoroughly discuss the multitude of options for array production, including surface capture strategies and array layouts (Uttamchandani and Yao, 2010). Typical SMMs vary in the amount of printed features with a range of 1000 to 16,000 printed features (Fig. 1). Nearly all SMM formats contain both positive and negative controls. Fluorescent dyes may be used as positive controls for SMM production, monitoring the fidelity of surface chemistry and washing steps, and they may serve as printed sentinels that help to frame the array. These dye sentinels are often helpful when analyzing the SMM image data, as they anchor the image analysis and array content files. The dyes are often chosen to be orthogonal from the dye used for detecting protein-small molecule interactions. For example, an amine-tagged version of fluorescein or Alexa Fluor 488 is often arrayed for the sentinel controls, and tagged antibodies or proteins are used for detection that is fluorescent in the red region (~650/670 nm). Other typical positive controls include small molecules with known protein partners that are commercially available such as biotin/streptavidin and rapamycin/FK506 (FKBP12). The positive controls can be useful in monitoring the quality of SMM production and are typically included in each of the printed subarrays that correspond to a given pin. If known ligands are available for the target of interest, these compounds may also be printed on the arrays if the compound is compatible with the chosen surface-capture strategy. For many recalcitrant targets, there are no known positive controls for the screen, as there are no known ligands. Depending on the nature of the surface and the chemical composition of the printed compounds, SMMs can typically be stored for 6 to 12 months at −20°C protected from light. This article focuses on the SMM screening procedure using HEK293 cell lysate expressing the target of choice

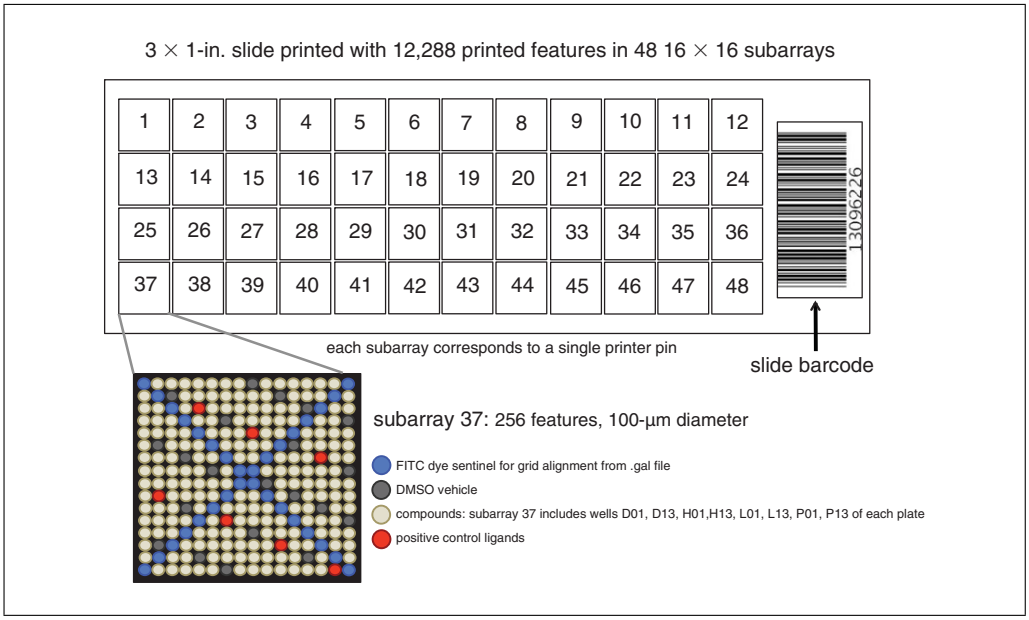


**Figure 1** Schematic diagram of a representative printed SMM. This particular array layout involves 12,288 features arrayed in 48 16 × 16 subarrays. Printed features are typically 150 µm in diameter for this array density. Each subarray contains 256 features and is created by a single pin from the 48-pin printhead. The subarrays are laid out in the same 4 × 12 configuration as the printhead. For subarray 37, the pin would pick up samples from wells D01, D13, H01, H13, L01, L13, P01, and P13 of each plate loaded onto the arrayer if printed in sequence. In this configuration, 32 features of each subarray are dedicated to printed dye sentinels for grid alignment (see data analysis section of the Basic Protocol). This configuration also features 20 DMSO vehicle controls and 8 known ligand-positive controls per subarray. Positive and negative controls can enable quality control analysis of the manufacture and screening processes as well as assist in data analysis methods. Barcodes are particularly useful when tracking slides, and the barcodes usually correspond to saved .tiff image files for screened slides and .gpr files corresponding to raw data analysis (see Basic Protocol section on Screening and Data Analysis).

tagged with HA tag (or any other appropriate tag, including encoded fluorescent tags such as GFP or mCherry). The preparation of the screening lysate is detailed in the Support Protocol.

## Alternate Strategy for Target Expression

### *Endogenous targets*

In some instances, a target can be sufficiently expressed endogenously to allow detection on the array. If highly specific antibodies are available for such proteins, it is preferable to screen in such conditions, thereby eliminating any non-physiologic interactions resulting from overexpression. Key considerations for this assay format include the quality of the antibody reagents and the possibility that direct antibody interactions with the protein may affect the interaction of a small molecule with the protein target (Bradner et al., 2006).

### *Exogenous targets*

For some targets, expression in HEK293T cells may not be the optimal cellular context, and therefore it may be more physiologically relevant to use a cell line that is more pertinent to the target in question. Ideally, the selected cell line grows well in cell culture conditions and is efficiently transfectable. If transient transfection is not suitable, a lentiviral infection can be used (*http://www.broadinstitute.org/rnai/public/resources/protocols*).

## SMALL-MOLECULE MICROARRAY (SMM) SCREENING

Many proteins are intrinsically disordered and are not amenable to purification procedures. This characteristic can impede small molecule screening campaigns aimed at identifying potential inhibitors. To overcome such critical problems in drug discovery, a protocol that uses HEK293 lysate expressing the desired target is described. In a lysate context, such targets tend to be more stable and their biological functions are largely preserved. These screening conditions are suitable for many types of proteins, in particular for the so-called un-druggable targets. It facilitates the screening of large small molecule collections and can lead to the discovery of putative direct binders, therefore significantly reducing the number of potential hits in subsequent secondary assays.

### *Materials*

Printed small-molecule microarray (SMM) slides, four replicate arrays (Uttamchandani and Yao, 2010; Casalena et al., 2012); store SMM slides at −20°C until ready
FKBP12 control protein, available in various tagged formats (e.g., His-tag, GST-tag, etc.) available from multiple suppliers (e.g., carrier-free and His-tagged, R&D Systems, cat. no. 3777-FK100)
TBST buffer (see recipe)
HEK293T lysate containing HA-tagged target protein (see Support Protocol)
MIPP lysis buffer (see recipe)
Mock lysate (see Support Protocol)
Anti-HA antibody, fluorophore-labeled (e.g., Cy5 or Cy3) (Cell Signaling, cat. no. 3444) secondary antibody.
Milli-Q water or equivalent
Rapamycin (LC Labs, cat. no. R-5000)
Alternative or additional controls:
    Alexa647-labeled streptavidin control protein (Life Technologies, cat. no. S21374)
    Biotin-cadaverine (Life Technologies, cat. no. A-1594)

4-well dishes (VWR, cat. no. 7321-424)
Rocking shakers, 4°C and room temperature
Laboratory Slide Spinner (Labnet, cat. no. C1303T)
GenePix 4200A microarray scanner (Molecular Devices)
GenePix Pro software (Molecular Devices)

### *Perform pilot experiments to optimize screening conditions*

To perform an optimal SMM lysate-based screen, several assay-development experiments must be designed and executed. Due to the varying nature of expression levels between cell lines and various target proteins, it is imperative that the assay development round be completed in the lysate of the expression systems chosen for the target. The goal of assay development is to optimize the ratio of protein target and total protein lysate concentrations to minimize background noise and maximize positive signals.

1. Before starting the SMM screen/assay development, place the required number of SMM slides at room temperature and allow them to equilibrate to ambient temperature.

    *Each assay development round should contain at least two identical slides per condition with duplicate compounds present on each slide, and should contain a positive control. For example, rapamycin, a small molecule that binds to FKBP12, and biotin, a small molecule that binds to streptavidin, may be printed throughout the array as controls.*

    *Stock concentrations and storage conditions for proteins can vary according to the nature of the protein. Protein stocks can be stored in nearly any buffer but caution should*

*be taken to avoid autofluorescent buffer components such as some detergents. If the autofluorescent potential of a given buffer composition is unknown, incubate a blank microscope slide in the desired buffer, dry, and scan at the desired detection wavelengths to determine if the buffer leaves a fluorescent film that will interfere with the assay.*

2. Prepare four solutions of purified FKBP12 protein diluted in 6 ml TBST at the following protein concentrations: 0.1 μg/ml, 0.25 μg/ml, 0.5 μg/ml, and 1 μg/ml.

   *One can consider using various tags as a control depending upon the detection strategy for the target of interest. The carrier-free and His-tagged sample from R&D Systems is recommended.*

3. Prepare another set of five FKBP12 protein solutions, except dilute the protein in cell lysate that has been diluted in MIPP buffer, as follows:

   a. Add 0.25 μg/ml FKBP12 to 6 ml diluted lysate containing 0.3 mg/ml total protein.

      *This concentration should yield very low assay signal.*

   b. Add 0.5 μg/ml FKBP12 to 6 ml diluted lysate containing 0.3 mg/ml total protein.
   c. Add 0.5 μg/ml FKBP12 to 6 ml diluted lysate containing 1 mg/ml total protein.
   d. Add 1 μg/ml of FKBP12 to 6 ml diluted lysate containing 0.3 mg/ml total protein.
   e. Add 1 μg/ml of FKBP12 to 6 ml diluted lysate containing 1 mg/ml total protein.

   *Total protein concentrations in stock lysate vary between 7 and 30+ mg/ml and should be quantified ahead of time using a suitable method such as BCA protein quantification.*

4. Place slides in 4-well dishes (4 slides/dish; five dishes required).

   *There should be a total of 18 slides, eight conditions from step 1 and ten conditions from step 2.*

5. Add 3 ml of each protein solution or diluted lysate solution to each slide.

6. Incubate dishes with slides covered, rocking gently for 2 hr at 4°C.

7. Remove covers of 4-well dishes, set aside, and fill the covers with ice-cold TBST (~5 ml each).

8. Remove the slides from 4-well dishes and place them in TBST-filled covers (4 slides/cover). Ensure that there is enough TBST in the covers to fully submerge slides. Gently rock for 5 min at room temperature.

9. Discard TBST, keeping slides in covers, and add an additional ~5 ml TBST. Wash slides a total of three times with 5 ml TBST each time.

10. While slides are washing, prepare a 1:1000 dilution of labeled anti-HA antibody into TBST buffer. Prepare a sufficient volume for the number of slides tested on that day assuming 3 ml volume per slide.

    *High-quality conjugated antibodies are available from a number of vendors. For example, a Cy5-labeled anti-HA antibody is available from Bioss Inc. (cat. no. bs-0966R-Cy5).*

11. Pipet the antibody solution into fresh 4-well dishes (3 ml per well). Remove slides from TBST and place them into antibody-filled wells.

12. Place dishes onto a shaker and gently rock for 30 min at room temperature.

13. Repeat steps 8 and 9 using Milli-Q water in place of TBST.

    *If anti-HA antibody is not labeled, use a secondary labeled antibody and repeat steps 8 through 13.*

14. Remove slides and place them in slide centrifugal spinners (minicentrifuge) for 30 sec.

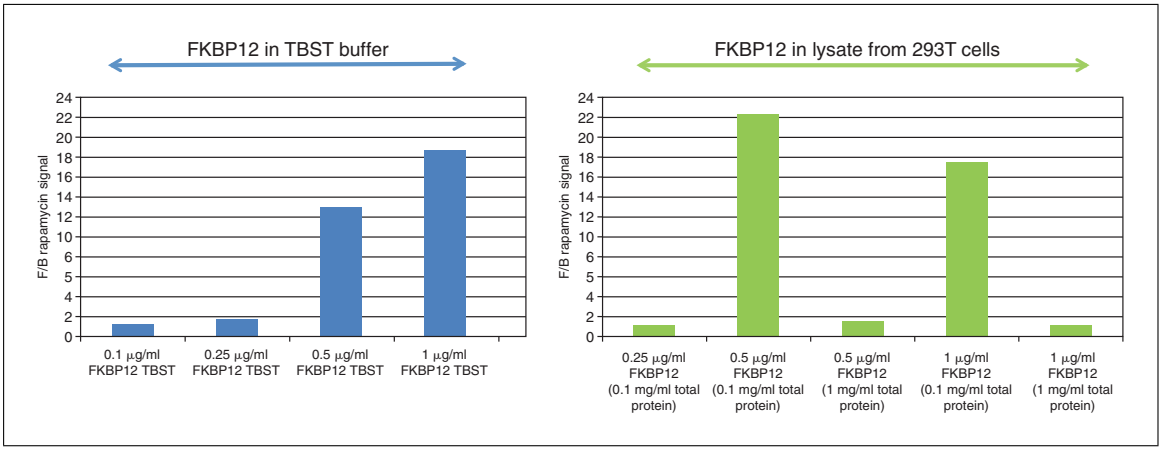


**Figure 2** FKBP12 binding to rapamycin as a representative protein-small molecule interaction on an SMM. Varying concentrations of pure FKBP12 were diluted in TBST (blue) or 293T lysates (green) and subjected to SMM screen using arrays containing rapamycin. Image analysis demonstrates that the signal intensity is significantly affected by the amount of FKBP12 protein screened. The *y*-axis represents the F/B ratios and the *x*-axis shows the FKBP12 protein concentrations in buffer or lysate, as indicated.

15. Place slides in GenePix scanner and scan at 635 wavelength, 600 PMT gain.

***Perform assay development data analysis***

16. Save scanned image slides as Tiff files.

17. Using a .gal file generated by the arrayer, align regions of interests (ROIs) over the printed spots using the preferred software.

    *A .gal file, otherwise known as gene-array list file from nucleic acid–based microarrays, is a file generated by the arrayer with the coordinates of each spot printed.*

    *GenePix 7.0 is the preferred software package that enables data analysis, joining assignment of array content to scanner output.*

18. After the ROIs are assigned, run image analysis and save the results file, often called a .gpr file.

    *Each saved analysis generates a .gpr and/or .txt file containing compound identifiers (IDs) and corresponding fluorescent signals. The .txt file is compatible with analysis in various software environments such as GenePix, SpotFire, and Excel, among others.*

    *Each spot will have an F635 value for the total fluorescence of that spot. There are four columns in the .gpr data file or .txt file that will be used for signal analysis: block column, row column, F635 median, and B635 median. The 'block' column and 'row' columns contain information about the spot location, the F635 median column contains the total fluorescence on the particular spot and the B635 median column contains fluorescence of the surrounding area around the spot.*

19. Using Excel or an equivalent program, generate F/B ratios by dividing F635 value by B635 value, which will result in a fold difference between a spot and its background.

    *This is the number that will be used as signal for a given spot.*

20. Focus the assay development analysis on rapamycin spot signals generated in most assay conditions tested.

    *As illustrated in Figure 2, the rapamycin signal will vary with FKBP12 concentrations and is also affected by the protein concentration in the lysate. In this example, it was determined that 0.1 and 0.25 μg/ml target protein are too low to yield a reliable signal, whereas 0.5 μg/ml or 1 μg/ml of target protein would be optimal if the lysate protein concentration was 0.1 mg/ml. FKBP12 binding to rapamycin is inhibited when the lysate protein is 1 mg/ml.*

### *Screen*

It is common for most HTS formats to first execute a small pilot screen. The purpose of the pilot screen is to validate or optimize the protein target/lysate concentrations mentioned above, aiming to minimize background noise and maximize positive signals. This part of the screening process uses four identical slides to test two different target protein concentrations in the lysate preparation. After the image analysis is performed, the hit rates in the two conditions are compared and the concentration with the optimal hit rate (<1%) will be used in the SMM screen. Perform the pilot screen as follows.

21. Prepare the following solutions:

    a. Solution 1: prepare 6 ml of 0.5 μg/ml target protein expressed in lysate containing 0.3 mg/ml total protein.
    b. Solution 2: prepare 6 ml of 1 μg/ml target protein expressed in lysate containing 0.3 mg/ml total protein.

22. Using concentration values obtained in Support Protocol, step 12, dilute stock lysate in MIPP buffer so the target protein concentration is ~0.5 μg/ml. Knowing the total protein concentration of stock lysate, calculate the total protein concentration after the above dilution. If it is <0.3 mg/ml, adjust diluted lysate concentration to 0.3 mg/ml using the mock HEK293 lysate.

23. Incubate, wash, and read-out two slides in solution 1 and two slides in solution 2 following steps 4 to 15.

24. Perform data analysis described in steps 25 to 29 and choose the ideal target protein/lysate concentration to be used in the SMM screen.

    *The ideal concentration setup is the one that produces the best signal-to-noise ratio for a given control.*

### *Perform screening and data analysis*

25. Prepare sufficient SMM-ready lysate (as determined in the pilot screen protocol), and incubate each SMM slide with 3 ml lysate solution according to the procedure described in steps 3 to 14.

26. Perform image analysis following steps 16 to 20.

    *Nearly all of the remaining data analysis steps can be carried out using simple statistical software such as EasyFit and simple operations are described in user-friendly guides that come with this software.*

27. Generate F/B distributions for every block of each slide screened, refer to slide layout from previous protocol (Clemons et al., 2010).

28. Fit the generated distributions to a Cauchy distribution model (Johnson et al., 1995) with F/B ratios <0.7 and >1.3 being excluded from the fitting.

    *A Cauchy distribution model, which is a continuous probability distribution, is characterized by its amplitude rather than its mean and variance, which are undefined. The major characteristic of a Cauchy distribution is a "heavy tail." Statistical software packages such as EasyFit can be used to fit the F/B distributions into a Cauchy model.*

29. After the model is fit on all the blocks, assign *P*-values for all data points, including the ones that were excluded. This step can also be done using the software program EasyFit. Compounds that show a *P*-value of ≤7% with an F/B ratio >1 in all four replicates are typically assigned as assay positives.

    *The number of unique assay positives divided by the total number of unique compounds screened constitutes the hit ratio, which is determined during the pilot screen. The ideal*

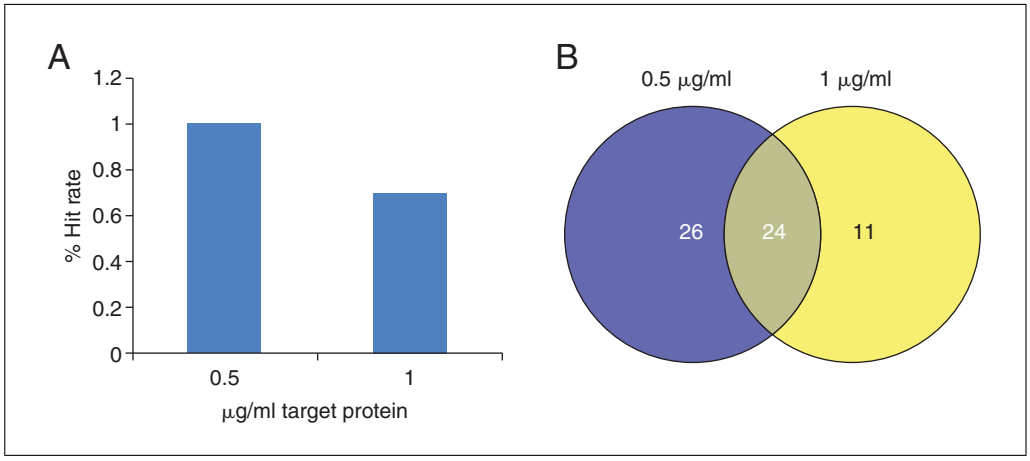


**Figure 3** Example of a pilot screen result. (**A**) The percent hit rate for each target protein concentration. (**B**) Venn diagram overlap for the hits identified in the pilot screen. Twenty-four hits were common between the two conditions, 26 hits were unique to the 0.5 μg/ml concentration, and 11 hits were unique to the 1 μg/ml concentration.

*hit rate should be <1%. In the example illustrated in Figure 3, both concentrations yielded an acceptable hit rate. The 0.5 μg/ml condition resulted in more hits than the 1 μg/ml condition, thus the 0.5 μg/ml concentration was used in the SMM screen.*

***SUPPORT PROTOCOL***

## PREPARATION AND QUANTIFICATION OF CELL LYSATE FOR SMALL MOLECULE MICROARRAY (SMM)

Certain target proteins are not amenable to purification procedures. Such proteins tend to be either unstable or nonfunctional when extracted from their physiological environments. To overcome this potential difficulty, a screening assay that uses cell lysates expressing the target of choice where physiological conditions (post-translational modifications, biological functions and folding) are maintained has been optimized. In this protocol, the target of interest is expressed in the HEK293T cell line.

### *Materials*

HEK293T cell line (ATCC# CRL-11268)
DMEM cell culture medium supplemented with 10% FBS and 1× penicillin/streptomycin
Protein target cloned in a mammalian expression vector (e.g., pcDNA3.1) driven by CMV promoter and harboring the HA tag at either N- or C- termini
OptiMEM medium (Life Technologies, cat. no. 11058-021)
Fugene6 transfection reagent (Promega, cat. no. E2691)
PBS buffer, ice cold
MIPP lysis buffer (ice-cold, see recipe) supplemented with protease inhibitors
BCA protein quantification reagents (Thermo Scientific, cat. no. 23225)
10 mg/ml bovine serum albumin (BSA) solution
HA-tagged purified protein (any protein similar in size with the desired target)
Anti HA antibody (Covance, cat. no. MMS-101P)
Appropriate labeled secondary antibody (i.e., for Odyssey detection or HRP-based)

Cell culture plates (10-cm diameter)
Cell scrapers
1.5-ml tubes
Eppendorf 5810 refrigerated centrifuge
96-well plate (non-sterile)
Spectrophotometer
Densitometry software (ImageQuant, ImageJ, Odyssey)

Additional reagents and equipment for western blot analysis (i.e., SDS protein gel, protein markers, buffers, transfer membrane, gel running chamber, power supply, gel transfer apparatus, etc.)

### *Transfect target*

1. One day prior to transfection, seed $2 \times 10^6$ HEK293T cells per 10-cm cell culture plate. Prepare several plates to generate mock lysates for use in assay development experiments aimed at optimizing the target concentration.

   *One $2 \times 10^6$ HEK293T cells per 10-cm plate can provide sufficient lysate to screen three SMM slides.*

2. Prepare transfection mixture according to the transfection reagent protocol: mix 5 μg plasmid DNA and OptiMEM medium with Fugene6 (or transfection reagent of choice) in a 1:5 ratio (according to manufacturer's instructions) and incubate 20 min at room temperature.

3. Transfer transfection mix onto cells containing fresh medium and incubate for 48 hr in a 37°C incubator.

### *Lyse cells and quantify total protein*

The SMM screening assay (see Basic Protocol) uses lysate generated from cells expressing the tagged target. Cell lysis buffer and its components permit extraction of the total cellular soluble protein and preserve the proteins including the target in a protected stable solution.

4. Wash plates containing the transfected cells one time with ice-cold PBS and remove the excess PBS buffer by maintaining the plates at an angle for 1 min.

5. Add 150 μl MIPP lysis buffer per plate and incubate 20 min on ice.

6. Collect lysates from plates using cell scrapers, transfer to a 1.5-ml tube, and centrifuge 10 min at 15,700 × *g*, 4°C. Transfer and save supernatant.

   *This is the stock cell lysate.*

7. Prepare BCA reagent mixture by mixing 20 μl reagent B in 980 μl reagent A and pipet 100 μl of this solution into a 96-well plate. Prepare triplicates for each concentration to be measured.

8. Generate a serial dilution for BSA concentrations (i.e., 10, 7.5, 5, and 2.5 mg/ml). Pipet 1 μl of each dilution and lysate into the triplicate wells, mix gently and incubate 30 min in the dark at room temperature.

9. Using a spectrophotometer, determine the absorption values at 562 nm and, based on the standard curve, calculate the total protein concentration of stock lysate.

### *Quantify target protein*

In the SMM screen, the target must be at a minimal concentration. A good starting point in this optimization is 0.5 μg/ml of diluted lysate (at ∼0.3 mg/ml total protein). This is easily achievable since in a typical overexpression experiment, the target is expressed at ∼0.5% of the total cellular protein.

10. Prepare SDS-PAGE samples containing target-expressing stock lysate at 25, 50, and 75 μg protein/sample. Prepare samples containing 0.1, 0.2, 0.4, and 0.8 μg purified HA-tagged protein.

    *The remaining lysate is the stock lysate and can be flash frozen in liquid nitrogen and stored for 24 to 48 hr at −80°C.*

11. Perform a western blot analysis for these samples using an anti-HA antibody coupled to a secondary readout (e.g., chemiluminescence or fluorescence) to visualize protein.

12. Use standard densitometry software (such as the software that often comes with an imager used to detect signal on blots) to quantify the signal in each lane. Based on the signal values for the purified reference, calculate how much target protein is present per milliliter of stock lysate. Use this information in Basic Protocol, step 22.

## REAGENTS AND SOLUTIONS

*Use Milli-Q purified water in all recipes and protocol steps.*

### *MIPP lysis buffer*

20 mM $NaH_2PO_4$, pH 7.2
25 mM β-glycerophosphate
2 mM EGTA
2 mM EDTA
0.5% (v/v) Triton X-100
Store up to 12 months at room temperature
Add fresh from 1 M stock solutions (stored at −20°C):
1 mM $Na_3VO_4$
5 mM NaF
1 mM DTT

### *TBS-T buffer*

50 mM Tris·Cl, pH 7.6
150 mM NaCl
Store up to 12 months at room temperature
Add fresh:
0.05% Tween 20
0.5% BSA (to antibody incubations)

## COMMENTARY

### Background Information

Small-molecule microarray screening is a high-throughput binding assay involving the fluorescence-based detection of protein-small molecule interactions where protein targets bind molecules arrayed on glass slides. The protein target used in the screening process can be in either purified format or from cell lysate. Lysates have advantages for use with SMM. Many protein targets are difficult to purify or unstable in standard buffer systems. Moreover, the presence of soluble cellular proteins in the lysates significantly reduces the background noise. The signal intensity is expressed as a function of signal-to-noise ratio (SNR). The final list of putative assay positives consists of compounds that score above a stringent arbitrary SNR cut-off and are reproducible in all replicates.

### Critical Parameters

#### *Choice of cellular context for target expression and target/lysate concentration*

In the authors' experience, target expression in HEK293T results in functional protein for most targets. For some targets, it may be more relevant for the target to be expressed in a specific cell type or cells subjected to a specific condition (e.g., heat shock or treatment with a small-molecule inhibitor). In such cases, additional optimization steps may be required to determine optimal levels of cell number, plasmid DNA for transient transfection, or lentiviral particles. It is also important to monitor the amount of expressed target and the total protein concentration. For example, high-total protein may decrease specific signal intensities, whereas low-total protein may result in an increase in false-positive hits.

### *Tag and antibody selection*

It is possible to use other epitope tags such as the Flag, His, SBP, and myc tags. However, it is essential that the anti-tag antibody be highly specific. In this protocol, the HA tag was selected due to the high specificity of the HA antibody.

### *Incubation steps*

It is important to minimize room temperature exposure during lysate and antibody incubation to avoid denaturation.

## Troubleshooting

### *No signals above background noise*

Hit rates in SMM screens can vary significantly and largely depend on the nature of the compound collection printed on the array as well as the nature of the protein target. In the case where no assay positives are observed during the pilot screen, a few steps and reagents should be verified and/or adjusted. For example, check that the scanning parameters are sufficiently sensitive for low signals. In addition, the amount of protein target may be increased.

### *Excessively high hit rate*

A great concern is a hit rate $>1\%$. This usually points to a potential surface chemistry problem often involving excessive surface activation or loading levels that results in high nonspecific binding. This issue must be addressed by the chemist charged with SMM manufacture.

## Anticipated Results

Based on previous experiences with various targets screened in lysate format, particularly in the case of transcription factors, it is expected that each screen will achieve hit rates in the range of 0.1% to 1%. For example, a recent screen that resulted in a published probe for the ETV1 oncogenic transcription factor yielded an average hit rate of $\sim$0.5% (Pop et al., 2014). Approximately 70% to 80% of SMM hits confirm binding in a secondary biophysical assay (e.g., thermal shift, SPR, ITC). Since SMM is a pure binding assay that does not involve interrogating biological functions, it is expected that the number of actual hits that bind and perturb a biological activity is significantly lower, especially for targets lacking known enzymatic activities that are easily followed in biochemical assays. Nevertheless, SMMs have successfully enabled the discovery of novel inhibitors (Hong et al., 2014).

## Time Considerations

The optimization of target expression in a particular cell line may take up to 1 week until the final optimal values are identified. A pilot screen, if necessary, will require an additional 3 to 4 days to establish the screening parameters. Once all conditions are known and all reagents are in hand, the screen itself can be performed in 1 day including image acquisition. If the compound collection is large ($>$100,000), an additional 1 or 2 days may be needed for image acquisition alone. Data analysis will require several days. In total, a target can be screened in 3 to 4 weeks when accounting for assay development, screening, and data analysis.

## Acknowledgements

The authors would like to thank Dylan V. Neel for editorial assistance and acknowledge funding from the National Cancer Institute (CA160860) that helped to support this work.

## Literature Cited

Bradner, J.E., McPherson, O.M., and Koehler, A.N. 2006. A method for the covalent capture and screening of diverse small molecules in a microarray format. *Nat. Protoc.* 1:2344-2352.

Casalena, D.E., Wassaf, D., and Koehler, A.N. 2012. Ligand discovery using small-molecule microarrays. *Methods Mol. Biol.* 803:249-263.

Clemons, P.A., Bodycombe, N.E., Carrinski, H.A., Wilson, J.A., Shamji, A.F., Wagner, B.K., Koehler, A.N., and Schreiber, S.L. 2010. Small molecules of different origins have distinct distributions of structural complexity that correlate with protein-binding profiles. *Proc. Natl. Acad. Sci. U.S.A.* 107:18787-18792.

Hong, J.A., Neel, D.V., Wassaf, D., Caballero, F., and Koehler, A.N. 2014. Recent discoveries and applications involving small-molecule microarrays. *Curr. Opin. Chem. Biol.* 18:21-28.

Johnson, N.L., Kotz, S., and Balakrishnan, N. 1995. Distributions. *In* Continuous Univariate Distributions, Vol. 2, pp. 362-432. Wiley, New York.

Miyazaki, I., Simizu, S., Okumura, H., Takagi, S., and Osada, H. 2010. A small-molecule inhibitor shows that pirin regulates migration of melanoma cells. *Nat. Chem. Biol.* 6:667-673.

Pop, M.S., Stransky, N., Garvie, C.W., Theurillat, J.P., Lewis, T.A., Zhong, C., Culyba, E.K., Lin, F., Daniels, D.S., Pagliarini, R., Ronco, L., Koehler, A.N., and Garraway, L.A. 2014. A small molecule that binds and inhibits the ETV1 transcription factor oncoprotein. *Mol. Cancer Therapeut.* 13:1492-1502.

Shi, H., Uttamchandani, M., and Yao, S.Q. 2010. A method for small molecule microarray-based screening for the rapid discovery of affinity-based probes. *Methods Mol. Biol.* 669:57-68.

Stanton, B.Z., Peng, L.F., Maloof, N., Nakai, K., Wang, X., Duffner, J.L., Taveras, K.M., Hyman,

J.M., Lee, S.W., Koehler, A.N., Chen, J.K., Fox, J.L., Mandinova, A., and Schreiber, S.L. 2009. A small molecule that binds Hedgehog and blocks its signaling in human cells. *Nat. Chem. Biol.* 5:154-156.

Uttamchandani, M. and Yao, S.Q. 2010. The expanding world of small molecule microarrays. *Methods Mol. Biol.* 669:1-15.

Uttamchandani, M., Walsh, D.P., Khersonsky, S.M., Huang, X., Yao, S.Q., and Chang, Y.T. 2004. Microarrays of tagged combinatorial triazine libraries in the discovery of small-molecule ligands of human IgG. *J. Comb. Chem.* 6:862-868.

Vegas, A.J. and Koehler, A.N. 2010. Detecting protein-small molecule interactions using fluorous small-molecule microarrays. *Methods Mol. Biol.* 669:43-55.

Walsh, D.P. and Chang, Y.T. 2004. Recent advances in small molecule microarrays: Applications and technology. *Comb. Chem. High Thr. Screen.* 7:557-564.

## Internet Resources

http://chembank.broadinstitute.org/

*A valuable resource is the Chembank database. It is a repository of existing SMM data sets performed for various types of targets and executed in either lysate format or with purified proteins. It can be especially beneficial in identifying and eliminating SMM hits with a promiscuous binding spectrum.*