



Published in final edited form as:

*Curr Protoc Chem Biol.* ; 6(4): 209–220. doi:10.1002/9780470559277.ch140101.

## Probing small molecule microarrays with tagged proteins in cell lysates

Marius S. Pop<sup>1,2,#</sup>, Dina Wassaf<sup>1,#</sup>, and Angela N. Koehler<sup>1,3</sup>

<sup>1</sup>Broad Institute of Harvard and MIT, Cambridge, Massachusetts

<sup>2</sup>The Dana-Farber Cancer Institute, Boston, Massachusetts

<sup>3</sup>Koch Institute for Integrative Cancer Research at MIT, Cambridge, Massachusetts

### Abstract

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 (SNR). 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.

### Keywords

small-molecule microarray; binding assay; cellular lysate; fluorescence; protein complex

### 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 (Shi et al., 2010; Uttamchandani et al., 2004; Uttamchandani and Yao, 2010; Vegas and Koehler, 2010; Walsh and Chang, 2004) and is not the focus of this protocol. The reader is referred to previous step-by-step protocols focused on the manufacture 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

<sup>#</sup>These authors contributed equally to this manuscript

#### Internet resources

A valuable resource is the ChEMBL database (<http://chembank.broadinstitute.org/>). 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.

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 can be analyzed using various methods but the most common analysis methods focus on signal-to-noise ratios.

The protocol presented herein describes a screening approach involving a tagged-target protein in the context of a cell lysate, and has been successfully applied previously to recalcitrant targets such as transcriptional regulators and extracellular factors (Miyazaki et al., 2010; Pop et al., 2014; Stanton et al., 2009). Lysate-based binding assays enable researchers to uncover 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 right 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 manufacture results in a linkage that is stable to cellular esterases prior to executing the screen. Please refer to the previously published protocols on SMM manufacture for lysate-compatible arrays (Casalena et al., 2012) or reviews that thoroughly discuss the multitude of options for array manufacture, including surface capture strategies and array layouts (Uttamchandani et al., 2010). Typical SMMs vary in the amount of printed features with a range of 1,000–16,000 printed features. Nearly all SMM formats contain both positive and negative controls. Fluorescent dyes may be used as positive controls for SMM manufacture, 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, we often array an amine-tagged version of fluorescein or Alexa Fluor 488 for the sentinel controls and we use tagged antibodies or proteins for detection that are 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 or FK506 (FKBP12). The positive controls can be useful in monitoring the quality of SMM manufacture 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 6–12 months at –20 °C and with protection from light. This article focuses on the SMM screening procedure using HEK293 cell lysate expressing the target of choice 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, therefore 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).

### 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.

## Support Protocols

### Support protocol A: Expression and target quantification

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 we have optimized a screening assay that uses cell lysates expressing the target of choice where physiological conditions (post-translational

modifications, biological functions and folding are maintained). In this protocol, the target of interest is expressed in HEK293 cell line.

### Materials

- HEK293T cell line (ATCC, cat# CRL-11268)
- The protein target cloned in a mammalian expression vector (i.e. pcDNA3.1) driven by CMV promoter and harboring the HA tag at either N- or C- termini
- Fugene6 transfection reagent (Promega cat# E2691)
- OptiMEM media (Life Technologies, cat# 11058-021) for transfection procedure
- DMEM cell culture media supplemented with 10% FBS and 1x penicillin/streptomycin
- Cell culture plates (10 cm diameter)
- Cell culture sterile cabinets and mammalian cell culture incubators

### Protocol

- 1 One day prior to transfection, seed 2 million HEK293 cells/10 cm plates. One plate can provide sufficient lysate to screen 3 SMM slides. Prepare several plates to generate mock lysates for use in assay development experiments aimed at optimizing the target concentration.
- 2 Prepare transfection mixture according to the transfection reagent protocol: mix 5 µg plasmid DNA and OptiMEM media with Fugene6 (or transfection reagent of choice) in a 1:5 ratio (according to manufacturer protocol) and incubate for 20 minutes at room temperature.
- 3 Transfer the transfection mix onto cells containing fresh media and incubate for 48 hours

### Support protocol B: Cell lysis and protein quantification

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

### Materials

- MIPP lysis buffer (ice cold) supplemented with protease inhibitors
- PBS buffer (ice-cold)
- BSA (bovine Serum Albumin) solution 10 mg/mL
- Cell scrapers
- Eppendorf 5810 centrifuge
- BCA protein quantification reagents (Thermo Scientific, cat # 23225)

- 96 well plate (non-sterile)
- Spectrophotometer

#### Protocol

- 4 Wash the plates containing the transfected cells with ice-cold PBS and remove the extra PBS buffer by maintaining the plates at an angle for one minute
- 5 Add 150  $\mu$ l lysis buffer per plate and incubate on ice for 20 minutes
- 6 Collect the lysates from plates using the cell scrapers, transfer to a 1.5 mL tube and centrifuge at 13000 rpm/4°C for 10 minutes. Transfer and save the supernatant. This is the stock cell lysate.
- 7 Prepare BCA reagent mixture by mixing 20  $\mu$ l reagent B in 1 mL reagent A and pipet 100  $\mu$ l of this solution in the 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  $\mu$ l of each dilution and lysate in the triplicate 96 wells mixing gently and incubate 30 min in the dark (i.e. drawer) at room temperature.
- 9 Using the spectrophotometer and determine the absorption values and based on the standard curve calculate the total protein concentration of stock lysate.

#### Support protocol C: Protein target quantification

In the SMM screen the target must be at a minimal concentration and a good starting point in this optimization is 0.5  $\mu$ g/ml of diluted lysate (approximately at 0.3 mg/mL). This is easily achievable since in a typical overexpression experiment the target is expressed at about 0.5% of the total cellular protein.

#### Materials

- HA-tagged purified protein (any protein similar in size with the desired target) to be used as a quantification reference
- Western blot materials and instruments (SDS protein gel, protein markers, buffers, transfer membrane, gel running chamber, power supply, gel transfer apparatus, etc)
- Anti HA antibody (Covance, cat# MMS-101P)
- Appropriate secondary labeled-antibody (i.e. for Odyssey detection or HRP based)
- Densitometry software (ImageQuant, ImageJ, Odyssey)

#### Protocol

- 10 Prepare SDS PAGE samples containing target-expressing stock lysate at 25, 50 and 75  $\mu$ g protein/sample. Prepare samples containing 0.1, 0.2, 0.4 and 0.8  $\mu$ g purified HA-tagged protein. The remaining lysate is the stock lysate and can be flash frozen in liquid nitrogen for short-term storage at  $-80^{\circ}\text{C}$ .

- 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 mL of stock lysate. Use this information for assay development procedure in step 21 of basic protocol.

## Basic Protocol: SMM Screening

### Materials

- Printed small-molecule microarray (SMM) slides, four replicate arrays (Casalena et al., 2012, Uttamchandani et al., 2010).
- HEK293T lysate containing HA-tagged target protein and mock lysate prepared as described in the support protocol
- MIPP lysis buffer (see recipe)
- Rocking shaker
- 4-well dishes for slide incubation from (VWR, 7321-424).
- HA antibody, fluorophore-labeled (e.g. Cy5 or Cy3) (Cell Signaling, cat# 3444) secondary antibody. 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 # bs-0966R-Cy5).
- TBS-T buffer (see recipe)
- Milli-Q water or equivalent
- Laboratory Slide Spinner (Labnet, C1303T)
- Microarray scanner GenePix 4200A (Molecular Devices)
- GenePix Pro software (Molecular Devices)
- FKBP12 control protein, this protein is available in various tagged formats (e.g. His-tag, GST-tag, etc.) from multiple vendors. One can consider using various tags as a control depending upon the detection strategy for your target of interest. The carrier-free and His-tagged sample from R&D Systems (cat # 3777-FK100) is recommended.
- Rapamycin, ligand of FKBP12 (LC Labs, R-5000)
- Alexa647-labeled Streptavidin control protein (Life Technologies, cat # S21374)
- biotin-cadaverine as ligand of streptavidin (Life Technologies, cat # A-1594)

## SMM screening procedure

In order to perform an optimal SMM lysate-based screen, a few assay development experiments need to 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 is 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 in order to minimize background noise and maximize positive signals when the assay is read out.

**A. Assay development procedure**—SMM slides are stored at  $-20^{\circ}\text{C}$  until executing the screening procedure. Before starting the SMM screen/assay development, acclimatize the slides to room 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. When in doubt about the autofluorescent potential of a given buffer composition, we typically suggest incubating a blank microscope slide in the desired buffer, drying and scanning at the desired detection wavelengths to determine if the buffer leaves a fluorescent film that will interfere with the assay.

- 1 Prepare four solutions of purified FKBP12 protein diluted in TBST (6 mL) with the following protein concentrations: 0.1  $\mu\text{g}/\text{mL}$ , 0.25  $\mu\text{g}/\text{mL}$ , 0.5  $\mu\text{g}/\text{mL}$ , and 1  $\mu\text{g}/\text{mL}$
- 2 Prepare another set of five FKBP12 protein solutions, only this time instead of diluting in TBST, dilute the protein in cell lysate that has been diluted in MIPP buffer (6 mL) as follows:
  - Add 0.25  $\mu\text{g}/\text{mL}$  FKBP12 to 6 mL of diluted lysate containing 0.3 mg/mL total lysate protein. This concentration should yield very low assay signal.
  - Add 0.5  $\mu\text{g}/\text{mL}$  FKBP12 into diluted lysate containing 0.3 and 1 mg/mL total lysate protein (6 mL each).
  - Add 1  $\mu\text{g}/\text{mL}$  of FKBP12 into diluted lysate containing 0.3 and 1 mg/mL total lysate protein (6 mL each), respectively.

Note: total protein concentrations in stock lysate vary between 7 and 30 mg/mL or greater and it should be quantified ahead of time using suitable method of choice such as BCA protein quantification.

- 3 Place slides in 4-well slide dishes (4 slides/dish; 5 dishes required).
- 4 Add 3 mL of each protein solution or diluted lysate solution to each slide. There should be a total of 18 slides, 8 conditions from step 1 and 10 conditions from step 2.

- 5 Incubate the dishes with slides at 4°C, rocking gently for two hours.
- 6 Remove covers of 4-well dishes, set aside and fill the covers with TBST (~5 mL each).
- 7 Remove the slides from 4-well dishes and place them in TBST-filled covers (4 slides/cover). Make sure there is enough TBST in the covers so slides are fully submerged. Place on gentle rocking for 5 min.
- 8 Discard TBST, keeping slides in covers, and add more TBST (~5 mL). Wash slides a total of 3 times with 5 mL TBST each time.
- 9 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.
- 10 Pipette the antibody solution into fresh 4-well dishes (3 mL per well). Remove slides from TBST and place them into antibody filled wells.
- 11 Place dishes on a gentle rocker at room temperature for 30 minutes.
- 12 Repeat steps 7 and 8 using Milli-Q water instead of TBST  
Note: If anti-HA antibody is not labeled, use a secondary labeled antibody and repeat steps 7 through 12
- 13 Remove slides and place them in slide centrifugal spinners for 30 seconds.
- 14 Place slides in Genepix scanner and scan at 635 wavelength, 600 PMT gain

## B. Assay development data analysis

- 15 Save scanned image slides as Tiff files.
- 16 Using a .gal file generated by the arrayer, align regions of interests (ROIs) over the printed spots using the preferred software. Genepix 7.0 is the preferred software package that enables data analysis joining assignment of array content to scanner output. *Note: 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.*
- 17 After the ROIs are assigned, run image analysis and save the results file, often called a “.gpr” file. Each saved analysis generates a .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.
- 18 Each spot will have an F635 value for the total fluorescence of that spot. There are 4 columns in the .gpr data 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 alternative 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 The assay development phase focuses on Rapamycin spots, which should generate signal in most of the 18 tested conditions.

*As illustrated in Figure 1, Rapamycin signal will vary with FKBP12 concentrations and is affected also by the protein concentration in the lysate. In this example it was determined that 0.1 and 0.25  $\mu\text{g}/\text{mL}$  target protein are too low to yield a reliable signal, whereas 0.5  $\mu\text{g}/\text{mL}$  or 1  $\mu\text{g}/\text{mL}$  of target protein would be optimal if the lysate protein concentration was 0.1  $\text{mg}/\text{mL}$ . FKBP12 binding to Rapamycin is inhibited when the lysate protein is at 1  $\text{mg}/\text{mL}$ .*

**C. Screening**—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 will use 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. The pilot screen is performed as follows:

- 21 Prepare the following solutions: Solution 1 (6 mL of 0.5  $\mu\text{g}/\text{mL}$  target protein expressed in lysate containing 0.3  $\text{mg}/\text{mL}$  total protein) and Solution 2 (6 mL of 1  $\mu\text{g}/\text{mL}$  target protein expressed in lysate containing 0.3  $\text{mg}/\text{mL}$  concentration). Using concentration values obtained in step 12 of **Support Protocol**, dilute stock lysate in MIPP buffer such that the target concentration is  $\sim 0.5 \mu\text{g}/\text{ml}$ . Knowing the total protein concentration of stock lysate, calculate the total protein concentration after the above dilution. If it is < 0.3  $\text{mg}/\text{ml}$  adjust diluted lysate concentration to 0.3  $\text{mg}/\text{ml}$  using the mock HEK293 lysate.
- 22 Incubate, wash and read-out two slides in Solution 1 and two slides in Solution 2 following instructions presented in the **Assay development procedure** (section A, steps 3–14)
- 23 Perform the data analysis procedure described in the **Screening and Data analysis** (Section D, steps 24–28) 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.

#### D. Screening and Data Analysis

- 24 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 the **Assay development** (Section A, steps 3–14).
- 25 Perform image analysis following the steps outlined in the **Assay development data analysis** (Section B, steps 1–6).

- 26 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. First, generate F/B distributions for every block of each slide screened (please refer to slide layout from previous protocol (Clemons et al., 2010).
- 27 Fit the generated distributions to a Cauchy distribution model (Johnson et al., 1994) 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.
- 28 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% or less with an F/B ratio >1 in all four replicates are typically assigned as assay positives.

Note: 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 hit rate should under 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.

### Reagents and solutions

Expression plasmid encoding the HA-tagged target protein

Transfection reagent: Use a highly efficient low toxicity reagent

Infectious lentiviral particles encoding the HA-tagged target protein

MIPP lysis buffer (use Milli-Q water or equivalent)

20 mM NaH<sub>2</sub>PO<sub>4</sub> [pH 7.2]

25 mM β-glycerophosphate

2 mM EGTA

2 mM EDTA

0.5% [v/v] Triton X-100

Add fresh from 1 M stock solutions (stored at –20C):

1 mM Na<sub>3</sub>VO<sub>4</sub>

5 mM NaF

1 mM DTT

TBS-T buffer

50 mM Tris [pH 7.6]

150 mM NaCl

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 our 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 is 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 so as 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 greater than 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 needs to be addressed by the chemist charged with SMM manufacture.

### Anticipated results

Based on our previous experience 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 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 about 0.5% (Pop et al., 2014). Approximately 70–80% 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 putative 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 a week until the final optimal values are identified. A pilot screen, if necessary, will require an additional 3–4 days to establish the screening parameters. Once all conditions are known and all reagents are in hand, the screen itself can be performed in one day including the image acquisition. If the compound collection is large (>100,000) an additional day or two 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.

### Acknowledgments

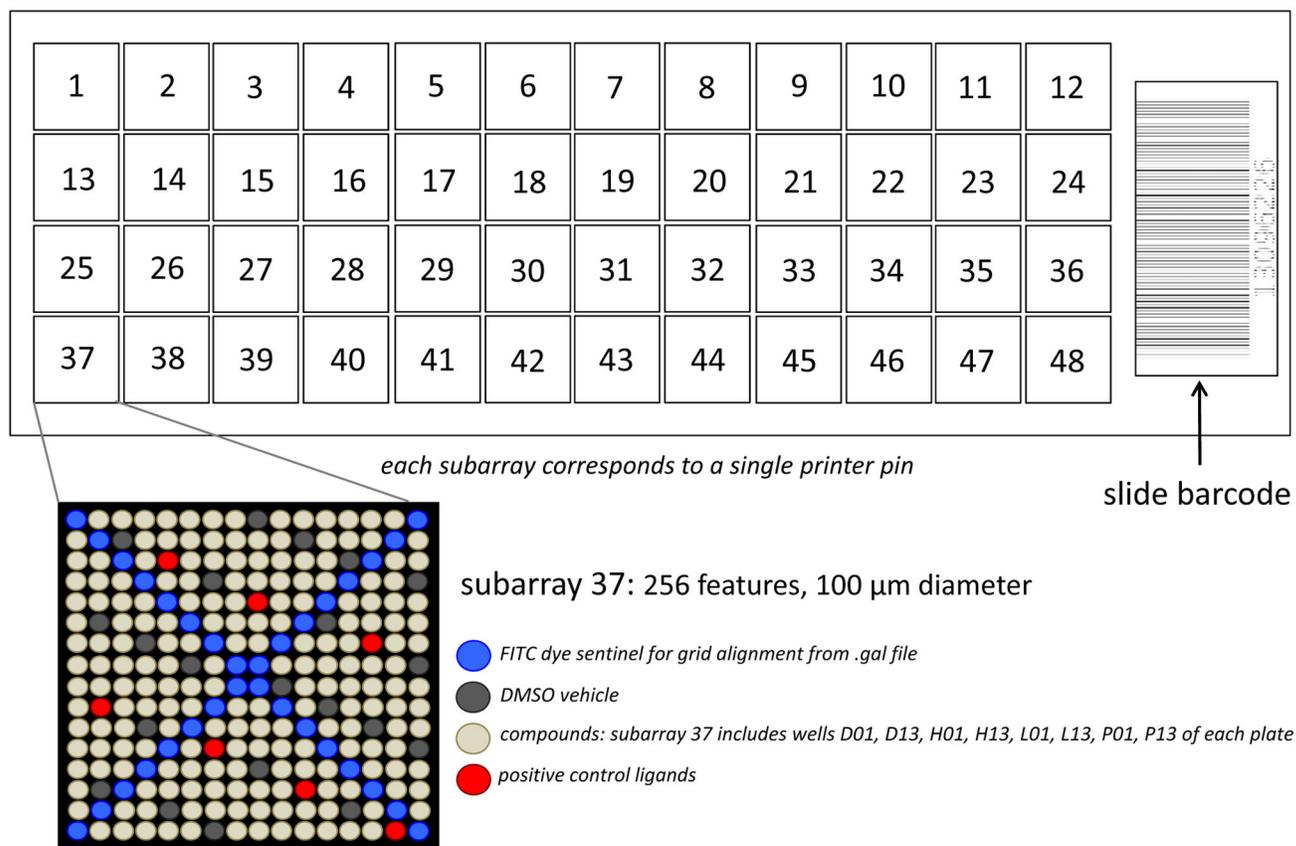
The authors would like to thank Dylan V. Neel for editorial assistance and acknowledge funding from the National Cancer Institute (CA160860) that helped to support this work.

### Literature Cited

- Bradner JE, McPherson OM, Koehler AN. A method for the covalent capture and screening of diverse small molecules in a microarray format. *Nature protocols*. 2006; 1:2344–2352.
- Casalena DE, Wassaf D, Koehler AN. Ligand discovery using small-molecule microarrays. *Methods in molecular biology*. 2012; 803:249–263. [PubMed: 22065230]
- Clemons PA, Bodycombe NE, Carrinski HA, Wilson JA, Shamji AF, Wagner BK, Koehler AN, Schreiber SL. Small molecules of different origins have distinct distributions of structural complexity that correlate with protein-binding profiles. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107:18787–18792. [PubMed: 20956335]
- Hong JA, Neel DV, Wassaf D, Caballero F, Koehler AN. Recent discoveries and applications involving small-molecule microarrays. *Current opinion in chemical biology*. 2014; 18:21–28. [PubMed: 24534749]
- Johnson, NL.; Kotz, S.; Balakrishnan, N. *Continuous Univariate Distributions*. Vol. 1. New York: Wiley; 1994. Chapter 16

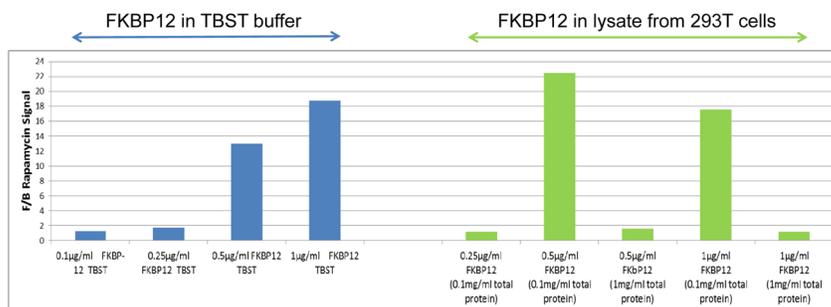
- Miyazaki I, Simizu S, Okumura H, Takagi S, Osada H. A small-molecule inhibitor shows that pirin regulates migration of melanoma cells. *Nature chemical biology*. 2010; 6:667–673.
- Pop MS, Stransky N, Garvie CW, Theurillat JP, Lewis TA, Zhong C, Culyba EK, Lin F, Daniels DS, Pagliarini R, Ronco L, Koehler AN, Garraway LA. A small molecule that binds and inhibits the ETV1 transcription factor oncoprotein. *Molecular cancer therapeutics*. 2014
- Shi H, Uttamchandani M, Yao SQ. A method for small molecule microarray-based screening for the rapid discovery of affinity-based probes. *Methods in molecular biology*. 2010; 669:57–68. [PubMed: 20857357]
- Stanton BZ, Peng LF, Maloof N, Nakai K, Wang X, Duffner JL, Taveras KM, Hyman JM, Lee SW, Koehler AN, Chen JK, Fox JL, Mandinova A, Schreiber SL. A small molecule that binds Hedgehog and blocks its signaling in human cells. *Nature chemical biology*. 2009; 5:154–156.
- Uttamchandani M, Walsh DP, Khersonsky SM, Huang X, Yao SQ, Chang YT. Microarrays of tagged combinatorial triazine libraries in the discovery of small-molecule ligands of human IgG. *Journal of combinatorial chemistry*. 2004; 6:862–868. [PubMed: 15530111]
- Uttamchandani M, Yao SQ. The expanding world of small molecule microarrays. *Methods in molecular biology*. 2010; 669:1–15. [PubMed: 20857353]
- Vegas AJ, Koehler AN. Detecting protein-small molecule interactions using fluororous small-molecule microarrays. *Methods in molecular biology*. 2010; 669:43–55. [PubMed: 20857356]
- Walsh DP, Chang YT. Recent advances in small molecule microarrays: applications and technology. *Combinatorial chemistry & high throughput screening*. 2004; 7:557–564. [PubMed: 15379627]

## 3x1 inch slide printed with 12,288 printed features in 48 16x16 subarrays



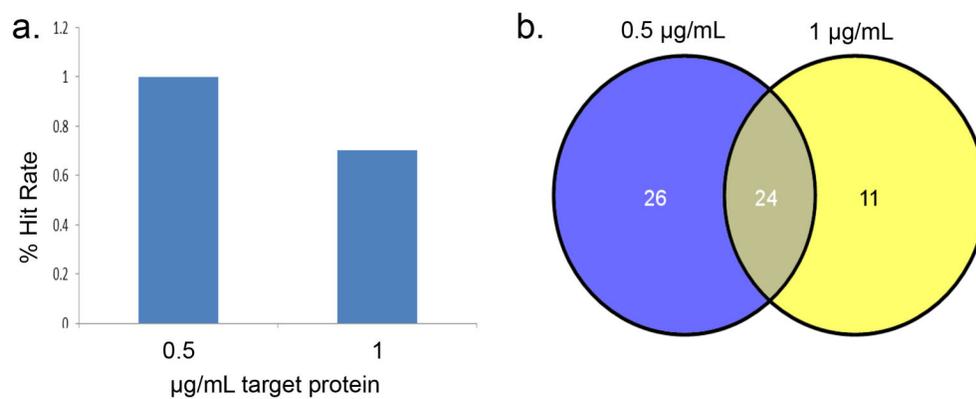
**Figure 1. Schematic diagram of a representative printed SMM**

This particular array layout involves 12,288 features arrayed in 48 16 $\times$ 16 subarrays. Printed features are typically 150  $\mu\text{m}$  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 $\times$ 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 plates 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 section on Data Analysis). This configuration also features 20 DMSO vehicle controls and 8 known ligand positive controls per subarray. Positive and negative controls can enable QC analysis of the manufacture and screening processes as well assist in data analysis methods. Barcodes are particularly useful when tracking slide and the barcodes usually correspond to saved .tiff image files for screened slides and .gpr files corresponding to raw data analysis (see section on Screening and Data Analysis).



**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 small molecule microarray 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.



**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. 24 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.