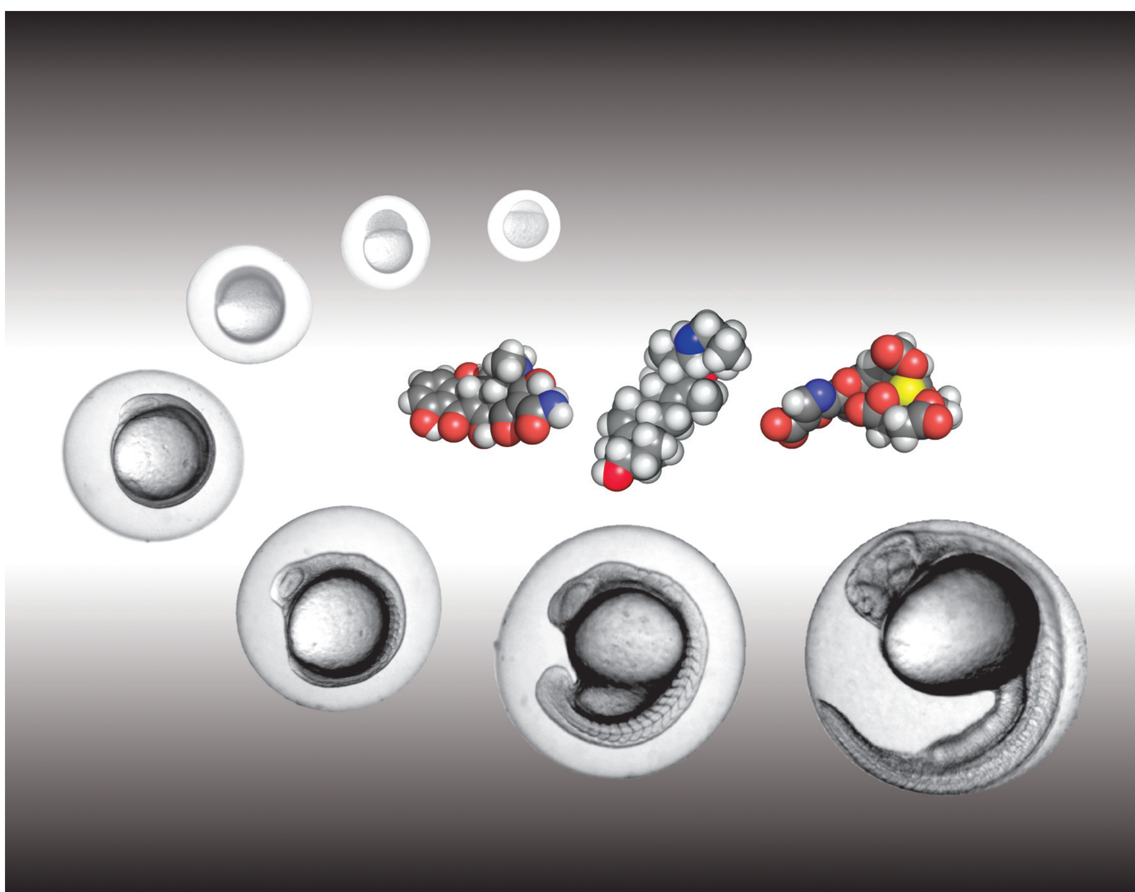
Chem Soc Rev

This article was published as part of the

2008 Chemistry–Biology Interface Issue

Reviewing research at the interface where chemistry meets biology

Please take a look at the full table of contents to access the other papers in this issue



Small-molecule microarrays as tools in ligand discovery†

Arturo J. Vegas,[a] Jason H. Fuller[b] and Angela N. Koehler*[b]

Received 21st April 2008
First published as an Advance Article on the web 20th May 2008
DOI: 10.1039/b703568n

Small molecules that bind and modulate specific protein targets are increasingly used as tools to decipher protein function in a cellular context. Identifying specific small-molecule probes for each protein in the proteome will require miniaturized assays that permit screening of large collections of compounds against large numbers of proteins in a highly parallel fashion. Simple and general binding assays involving small-molecule microarrays can be used to identify probes for nearly any protein in the proteome. The assay may be used to identify ligands for proteins in the absence of knowledge about structure or function. In this *tutorial review*, we introduce small-molecule microarrays (SMMs) as tools for ligand discovery; discuss methods for manufacturing SMMs, including both non-covalent and covalent attachment strategies; and provide examples of ligand discovery involving SMMs.

1. Introduction

Deciphering the information encoded in the functional genome, which includes thousands of characterized proteins and uncharacterized predicted gene products, is a key challenge for researchers in the post-genomic era. While the number of human protein-coding genes has been estimated to be almost 21 000,[1] processes such as alternative splicing[2,3] and post-translational modification[4] give rise to an expanded and dynamic functional proteome. It has been estimated that at least 10^6 biomolecules are required to maintain the integrity of human cells.[4,5] Understanding the roles of these biomolecules within complex, living systems requires new tools that can be used to explore the functional genome on a large scale. Complementary and high-throughput approaches rooted in comparative genomics,[6] proteomics,[7] and chemical biology[8] can be used to expedite our understanding, offering clues to biological function as well as validating disease targets for therapeutic intervention.

[a] *Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA*
[b] *Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. E-mail: koehler@fas.harvard.edu; Fax: +1-617-324-9601; Tel: +1-617-324-9634*

† Part of a thematic issue examining the interface of chemistry with biology.

Increasingly, small molecules are used directly as tools to study functions of proteins and cellular processes. Representative uses for tool compounds in chemical biology have been reviewed elsewhere and include enzyme inhibitors, receptor modulators, modulators of protein–protein interactions, aids in protein crystallization, detection or imaging agents, and activity-based probes.[8,9] Whereas genetic approaches can be applied generally to study nearly any protein target, chemical approaches are currently limited by the fact that relatively few proteins have known small-molecule partners. Ideally, small-molecule activators and inhibitors of any given function would be available to researchers interested in studying any given protein in the proteome.[8] Design of such small molecules in the absence of structural information about each protein in the proteome is a daunting task. Therefore, high-throughput, cost-effective, and general ligand discovery methods that require little or no information about protein structure or function are needed in order to identify specific small-molecule probes of each protein function.

Several high-throughput screening (HTS) approaches to ligand discovery have been described and reviewed elsewhere.[8–10] Two main HTS approaches, often referred to as forward and reverse chemical genetics in analogy to forward and reverse genetic screens, offer complementary paths to identifying bioactive molecules.[9] Cell-based, phenotypic

Arturo Vegas received a BA in Biology from Cornell University in 2001 and a PhD in Chemistry from Harvard University in 2008. His doctoral studies focused on developing methods to identify inhibitors and ligands of histone deacetylases and were carried out under the direction of Stuart Schreiber. He is currently a postdoctoral fellow working in the laboratory of Robert Langer at MIT where his research focuses on developing novel polymers for use in drug delivery.

Jason Fuller received a BS in Biotechnology from Worcester Polytechnic Institute in 2006. Prior to joining the Broad Institute as a research associate, he worked in the pharmaceutical industry and focused on developing methods for combinatorial chemistry. His current research involves the use of microarray-based binding assays to identify allosteric inhibitors of histone deacetylases. He plans to pursue doctoral studies in the area of synthetic biology.

Angela Koehler received a BA in Biochemistry and Molecular Biology from Reed College and studied with Stuart Schreiber at Harvard University for her PhD in Chemistry, which she received in 2003. Upon graduation, she became an Institute Fellow at the Broad Institute of Harvard and MIT. Her research interests focus on developing general ligand discovery methods and applying them to protein targets that are considered difficult to ligate, such as transcription factors.

This journal is © The Royal Society of Chemistry 2008

screens involve scoring small molecules based on their ability to alter cellular phenotypes (*e.g.* cell viability, transcriptional activity, nuclear translocation of a protein, or altered morphology). This forward chemical genetic approach allows a researcher to directly screen for compounds that induce a phenotype of interest in a cellular context but also requires additional investigation to identify the protein target(s) responsible for the observed phenotype. Unfortunately, a systematic and general approach to target identification remains elusive.[11] The reverse chemical genetic screening approach involves functional small-molecule assays targeted directly to a protein of interest (*e.g.* enzymatic assay or protein–DNA interaction studies). These assays are typically performed in isolation from the cell. Once a compound that targets a given protein is identified, the challenge is to identify whether or not the small molecule has an effect in a cellular context. While target-based screens directly identify compounds that modulate a protein of interest, there is no guarantee that the compounds will be cell-permeable or that they will affect the protein in a way that results in a functional phenotypic outcome in the cell. Broad phenotypic studies are usually required to understand the consequences of modulating the target. From the perspective of drug discovery, targeting a specific protein may or may not have a desired therapeutic consequence and considerable effort is invested validating target proteins prior to executing full-scale HTS campaigns. Both HTS approaches have successfully delivered tool compounds that may be useful in validating specific protein targets for therapeutic design.[8,9]

Unfortunately, the majority of targeted or phenotypic HTS screens require some advance knowledge of protein structure and function or a considerable amount of assay development to execute. Simple, high-throughput, and general binding screens may provide a cost-effective alternative to conventional phenotypic or functional target-based HTS.[12] In this scheme, compounds are selected simply based on their ability to ligate a protein of interest in a high-throughput primary assay. The positives are then evaluated in downstream functional or phenotypic assays. Investment in functional or phenotypic assay development occurs only when ligands are in hand. Simple binding screens may identify ligands with the ability to activate or inhibit a given protein function. A single binding screen may identify compounds that act on a protein target in different modes (*e.g.* active-site and allosteric modulators). Ideally the compounds are subjected to many parallel proteins in an effort to evaluate binding specificity. Examples of techniques for detecting binding interactions between proteins and small molecules have been reviewed elsewhere and include fluorescence polarization, isothermal titration calorimetry, fluorescence-based thermal shifts, surface plasmon resonance, three-hybrid assays, and on-bead binding assays.[12,13] Many of these methods are limited by throughput, quantity of protein or small molecule required for the assay, or the need for assay development based on prior structure of the protein. In this tutorial review, we introduce small-molecule microarrays (SMMs) as an alternative general binding assay, compatible with nearly any type of protein without advanced knowledge of structure or function, and present examples of ligand discovery projects using SMMs.

Taking a cue from the world of DNA microarrays and whole genome expression profiling, Schreiber and co-workers reported the fabrication of microarrays containing small molecules for use in detecting interactions with proteins.[14] Since the initial publication, a number of notable improvements in both SMM manufacturing and screening have been reported by several investigators and reviewed elsewhere.[12,15,16] Typically, nanolitre volumes of small molecules are arrayed onto functionalized glass microscope slides (25 mm × 75 mm) and immobilized using one of several covalent or non-covalent methods (Fig. 1). Depending on the immobilization chemistry and the type of small molecule printed, microarray features vary in diameter from 50–300 μm. Microarrays containing

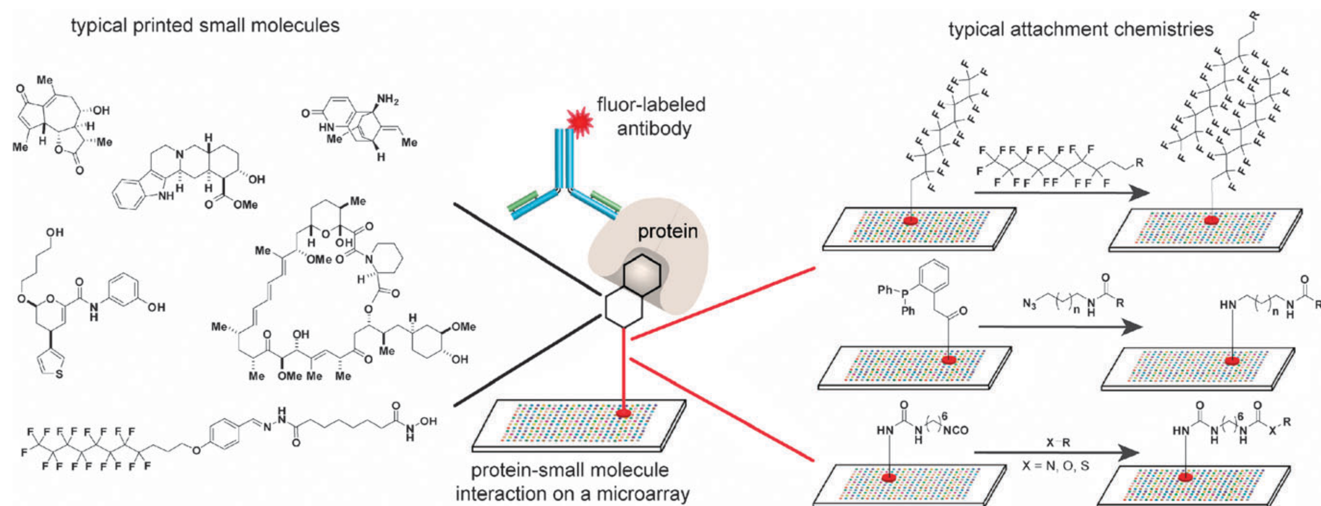
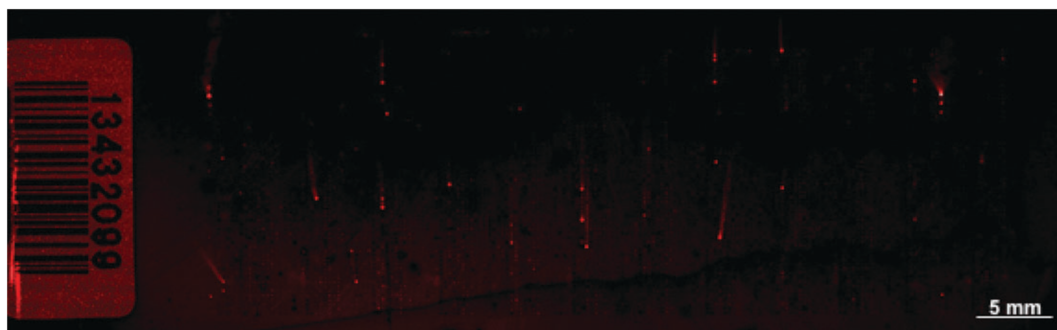


Fig. 1 Preparation and screening of small-molecule microarrays (SMMs). Small molecules from a variety of sources, including natural products, bioactives, commercial compound collections, and products of diversity-oriented syntheses (left), are arrayed onto chemically modified glass microscope slides. Compounds are attached to the surface using any one of a number of covalent or non-covalent methods (right). Printed arrays are incubated with a protein of interest and binding is often detected using a fluorescently labeled antibody against the protein itself or an epitope tag (center).

This journal is © The Royal Society of Chemistry 2008



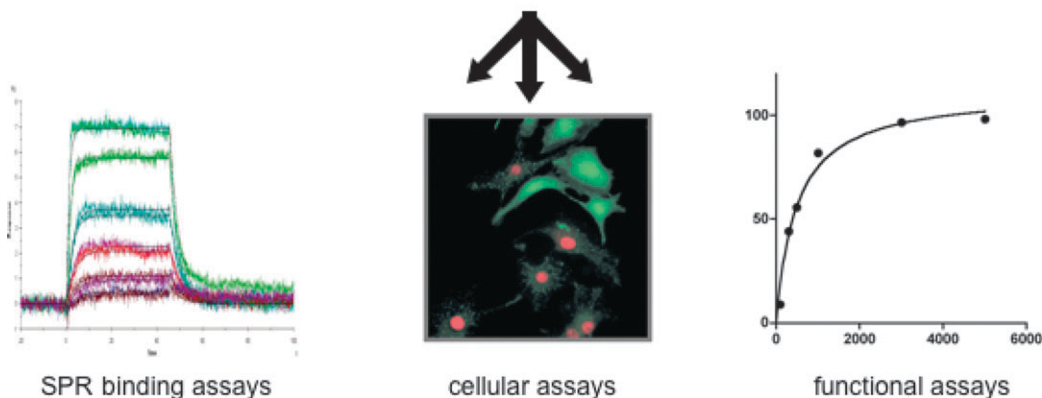


Fig. 2 Screened slides are scanned for fluorescence using a standard microarray scanner. Putative binders may be evaluated in secondary binding assays (*e.g.* surface plasmon resonance, thermal shift, calorimetry), cell-based phenotypic assays, or functional biochemical assays.

nearly 11 000 different small molecules have been reported and illustrate the miniaturized nature of the assay.[12,13] Published reports have described SMMs containing small molecules from a variety of sources including products of diversity-oriented syntheses, combinatorial libraries biased toward specific enzymatic activities, libraries of peptidomimetics, carbohydrates, known bioactive compounds, commercial compound collections, purified natural products, natural product extracts, and FDA-approved drugs.[12,17–19]

The microarrays are incubated with a protein of interest and interactions are usually detected using a fluorescence-based readout with a standard microarray scanner or by surface plasmon resonance (SPR).[12,20] When a fluorescence-based readout is used, purified proteins are often detected *via* a labeled antibody against an epitope tag or an antibody directly against the protein (Fig. 1). SMMs are also compatible with screens involving cell lysates containing endogenous proteins using antibodies against the protein target,[18] or overexpressed, epitope-tagged proteins.[12,19] Expressable and fluorescent tags, such as green fluorescent protein (GFP), have also been used successfully in screens involving lysates.[19] The ability to screen directly from cell lysates can save time and effort by bypassing the purification process. Proteins obtained from cellular lysates may be more likely to retain proper conformation for activity or contain relevant post-translational modifications. Lysate screens provide an opportunity to screen proteins that reside and act in protein complexes; the direct interaction of the immobilized small molecule may or may not involve the protein of interest, but rather, a direct interaction between the small molecule and another complex member. Secondary binding assays (*e.g.* SPR, fluorescence polarization, isothermal calorimetry, thermal shifts) may be used to determine whether or not the small molecule binds directly to the target protein or to another protein in the complex. In some cases, additional target identification studies, similar to those required for cell-based, phenotypic screens, may be required to determine which complex member interacts directly with the small molecule.

After SMM screening positives have been identified, compounds are evaluated in secondary binding assays, functional assays, or phenotypic assays (Fig. 2). In one report, SPR (Biacore) was used to evaluate more than 100 protein–small molecule interactions discovered using SMMs in secondary assays.[19] In these studies, the proteins were immobilized to dextran-coated sensor surfaces and small molecules were injected in solution at varying concentrations. Of the interactions tested, 86% retested as binders with dissociation constants of 0.5–20 μM. Several ligands with varying affinities have been discovered using this approach and are reviewed elsewhere.[9,10,12,15] Among others, representative proteins with ligands successfully identified using this approach include kinases,[21,22] proteases,[23] and transcriptional regulators.[18,24,25] Examples of screens involving these protein classes will be presented in this tutorial review.

2. Manufacturing small-molecule microarrays

The first step of any SMM experiment involves design and fabrication of the chips containing probe molecules of interest. Immobilization methods must take both orientation of display and molecular stability into account. Most types of SMMs are

Table 1 Summary of common attachment methods for preparing SMMs

Attachment method	Surface	Coupling partners	References
Michael addition	Maleimide	Thiol	14
Silyl ether formation	Silyl chloride	Primary alcohol	24–26
Oxime formation	Glyoxylyl	Aminoxyl	27
Thiazolidine ring formation	Glyoxylyl	1,2-Amino thiol	27
1,3-Dipolar cycloaddition	Terminal alkyne	Azide	28, 29
Diels–Alder	Benzoquinone	Cyclopentadiene	30
Capture of heteroatoms with acidic proton	Diazobenzylidene	Phenol, carboxylic acid	31
Staudinger ligation	Phosphane	Azide	32
Amide formation	Activated ester	Amine	33
Epoxide opening by hydrazide	Epoxide	Hydrazide	34, 35
Photolithography	Various	Various	36
Photoactivated crosslinking	Diazarine, aryl azide	Various	20, 37, 38
Non-selective isocyanate capture	Isocyanate	Various	13, 17, 19, 21, 41
Self-sorting capture *via* hybridization	Oligonucleotides	PNA	23
Fluorous affinity capture	C_8F_{17}	C_8F_{17}	18, 39, 40
On-array enzymatic synthesis	Various	Various	22

prepared by immobilizing the small molecules on chemically treated glass microscope slides using either microcontact spotters or piezoelectric deposition. Microlithography, using either masks or optical methods, is another common method for fabricating microarrays and is used routinely to prepare oligonucleotide arrays.[9,15] Most reports of SMM manufacturing involve microcontact printing. A number of capture strategies have been developed for microarrays and include both covalent attachment of the probes to the surface and non-covalent deposition (Table 1).[13–42] A detailed review of several successful covalent and non-covalent attachment chemistries has been published elsewhere.[15]

2.1 Covalent immobilization

Most of these approaches involve mild and selective coupling reactions. The first SMMs were prepared using the Michael addition reaction involving molecules containing free thiols printed onto slides coated with vinyl sulfone or maleimide

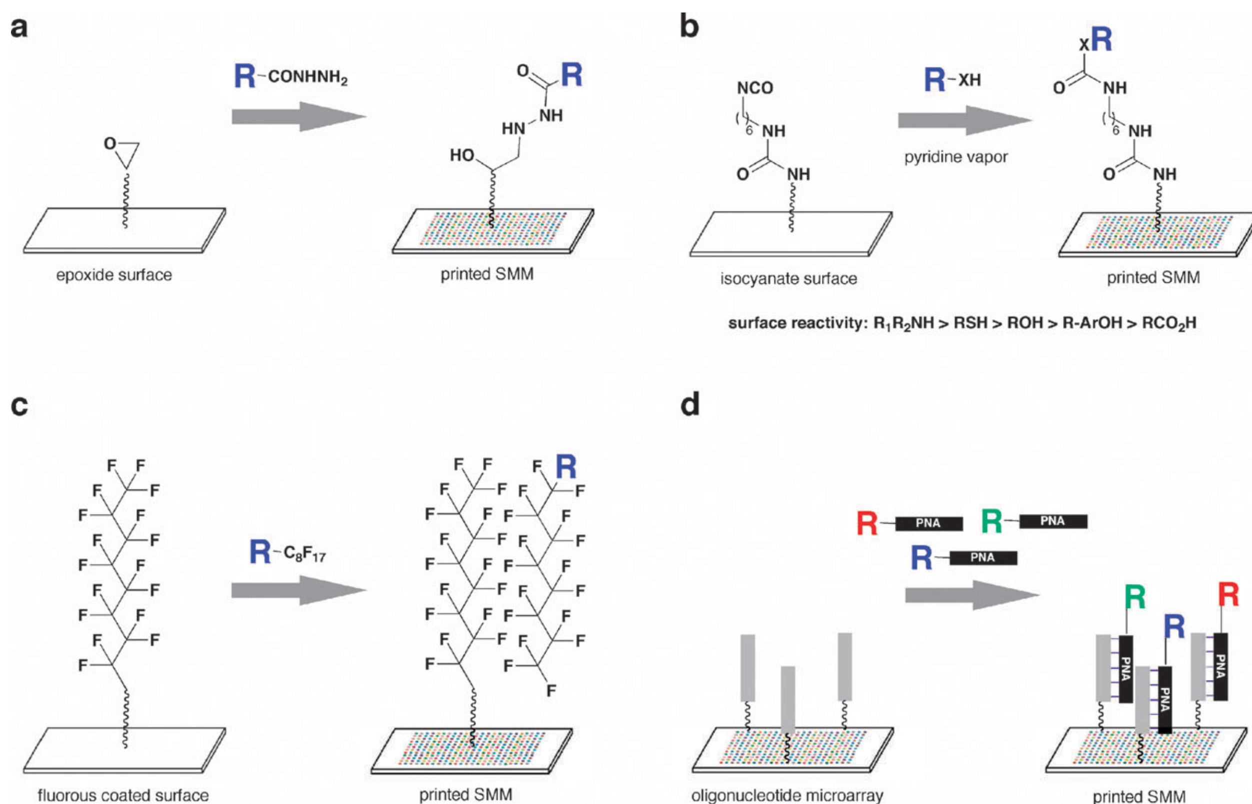


Fig. 3 Representative approaches to immobilizing small molecules. (a) Printing hydrazides on epoxide-coated surfaces is an example of selective covalent immobilization. The epoxide will react with the hydrazide selectively over other nucleophilic functional groups. (b) Isocyanate-coated surfaces react with a variety of nucleophilic functional groups and can be used to covalently immobilize small molecules in a non-selective fashion. (c) Fluoroalkylsilane-coated slides may be used to non-covalently capture polyfluorocarbon-tagged small molecules *via* a fluorous affinity interaction. (d) Libraries of PNA-encoded small molecules screened in solution may be captured selectively onto arrays *via* hybridization on oligonucleotide microarrays.

 This journal is © The Royal Society of Chemistry 2008

groups.[14] While this strategy proved successful it was not general as most compound screening collections do not contain a high proportion of free thiols. Many more compounds coming from both combinatorial libraries and natural product collections contain amino and hydroxy groups. While these compounds may be coupled to carboxy-modified glass *via* amide or ester bond formation, Chang and co-workers chose to print a library of 2688 amine-containing molecules on slides coated with *N*-hydroxysuccinimide (NHS) activated esters.[33] This approach obviates the need for additional coupling catalysts. Surfaces treated with silyl chloride have been used to capture 12 396 compounds from diversity-oriented synthesis (DOS) that contain primary alcohols.[24–26] A DOS library of 6336 compounds containing aryl alcohols was captured on diazobenzylidene surfaces.[31] Readily commercially available epoxide-coated glass slides have been used to capture hydrazide-tagged small molecules (Fig. 3a) and carbohydrates.[34,35] Several other mild and chemoselective immobilization strategies have been described and include immobilization through formation of oximes, hydrazones, or thiazolidine ring formation on glyoxylyl surfaces,[27] 1,3-dipolar cycloaddition,[28,29] Diels–Alder reaction,[30] and Staudinger ligation.[32] Most of the surface capture methods take advantage of a reactive functional group that is introduced as part of their synthesis and biases the orientation of the small molecule on the surface. Motivated by the need to increase molecular diversity on SMMs, non-selective approaches to capturing compounds have been adopted. An isocyanate-mediated capture strategy (Fig. 3b) was used to print nearly 10 000 known bioactive small molecule natural products, and small molecules originating from several diversity-oriented syntheses.[13,19] Isocyanates react with a variety of nucleophilic functional groups thereby increasing the number of small molecules, from natural or synthetic sources, that may be printed on a single surface. This approach was recently extended by Schmitz *et al.* to capture natural product extracts of varying degrees of purity and detect protein–small molecule interactions on extract microarrays.[17] Kanoh *et al.* prepared microarrays of approximately 2000 natural products and drugs by photocrosslinking compounds on trifluoromethylaryldiazarine-coated surfaces.[37] Using this approach, photogenerated carbenes react with the printed compounds in manner that is independent of functional group. Pei *et al.* prepared carbohydrate microarrays using a double photoligation strategy involving perfluorophenylazides (PFPAs).[38] PFPA-derivatized carbohydrates were immobilized on a poly(ethylene oxide) surface by photo-initiated insertion. Both the isocyanate and photo-crosslinking strategies present the possibility of printed compounds occupying multiple modes of orientation within a given spot, effectively increasing the number of binding modes that a given probe protein can sample. Other covalent approaches to SMM manufacture include on-array synthesis. For example, Kodadek and co-workers used photolithography to synthesize arrays containing cyclic peptides and added side chains to the pre-synthesized cyclic cores in a combinatorial fashion.[36] Dordick and co-workers used *in vitro* metabolic pathway construction to synthesize natural product analogues directly on the microarrays and identified three inhibitors of Fyn tyrosine kinase.[22]

2.2 Non-covalent immobilization

Non-covalent methods have traditionally been used to make microarrays of nucleic acids and proteins.[42] Slides coated with aminosilane or poly-L-lysine have been used to randomly capture oligonucleotides, proteins, and cells *via* electrostatic interactions or passive adsorption. Similarly, nitrocellulose has been used as a substrate for capture of DNA, proteins, and carbohydrates. Biomolecules can also be biotinylated and printed on streptavidin-coated surfaces. Microwells and microdroplets have been adapted to the microarray format in an effort to carry out experiments in solution.[43,44] More recently, non-covalent immobilization approaches have been developed for fabricating SMMs. Winssinger and co-workers prepared SMMs containing a PNA-encoded tetrapeptide acrylate library *via* sequence-specific hybridization to an oligonucleotide microarray (Fig. 3d).[23] The small molecules covalently linked to a PNA that non-covalently anchors the molecule to the array solid support. This approach involves encoding combinatorial libraries that may be screened in solution and subsequently immobilized and decoded *via* self-sorting. Taking advantage of the highly specific fluorous affinity interaction, Pohl and co-workers non-covalently captured polyfluorocarbon-tagged carbohydrates on fluoroalkylsilane-coated slides (Fig. 3c) and validated the fluorous-based microarrays as a screening tool for carbohydrate-binding proteins.[39] The authors identified a fluoroalkylsilane-coated glass surface with low intrinsic fluorescence and printed fluorous-tagged carbohydrates onto the surface, yielding printed features with reasonable spot diameters and morphology. The fluorous-based carbohydrate microarrays were compatible with incubation of protein in aqueous buffers without appreciable diffusion of the spotted carbohydrates. More recently, Spring and co-workers validated the use of fluorous-based microarrays for drug-like small-molecules by exploring known protein–small molecule interactions.[40] The authors observed excellent signal-to-noise ratios with fluorescently labelled protein, consistent with the results from Pohl and co-workers. More recently, fluorous-based microarrays were used to print and detect other types of small molecules. Schreiber and co-workers recently used the fluorous approach to print collections of fluorous-tagged compounds that are structurally biased to bind and inhibit histone deacetylases.[18]

3. Ligand discovery involving SMMs

Protein–ligand interactions of varying affinities have been discovered using SMMs. Several proteins from different functional classes including transcription factors, immunoglobulins, proteases, and kinases, have been successfully targeted with small molecules discovered using the SMM approach (Fig. 4). For example, haptamide and uretupamine bind and modulate two yeast proteins, Hap3p and Ure2p, respectively, involved in transcriptional regulation and nutrient-sensing.[24,25] Several ligands for calmodulin with significantly different core molecular scaffolds have been identified using SMMs. For example, calmodioxane and calmoduphilin are products of diversity-oriented syntheses.[12,31] NPC-15437, a known inhibitor of protein kinase C, binds to calmodulin

This journal is © The Royal Society of Chemistry 2008

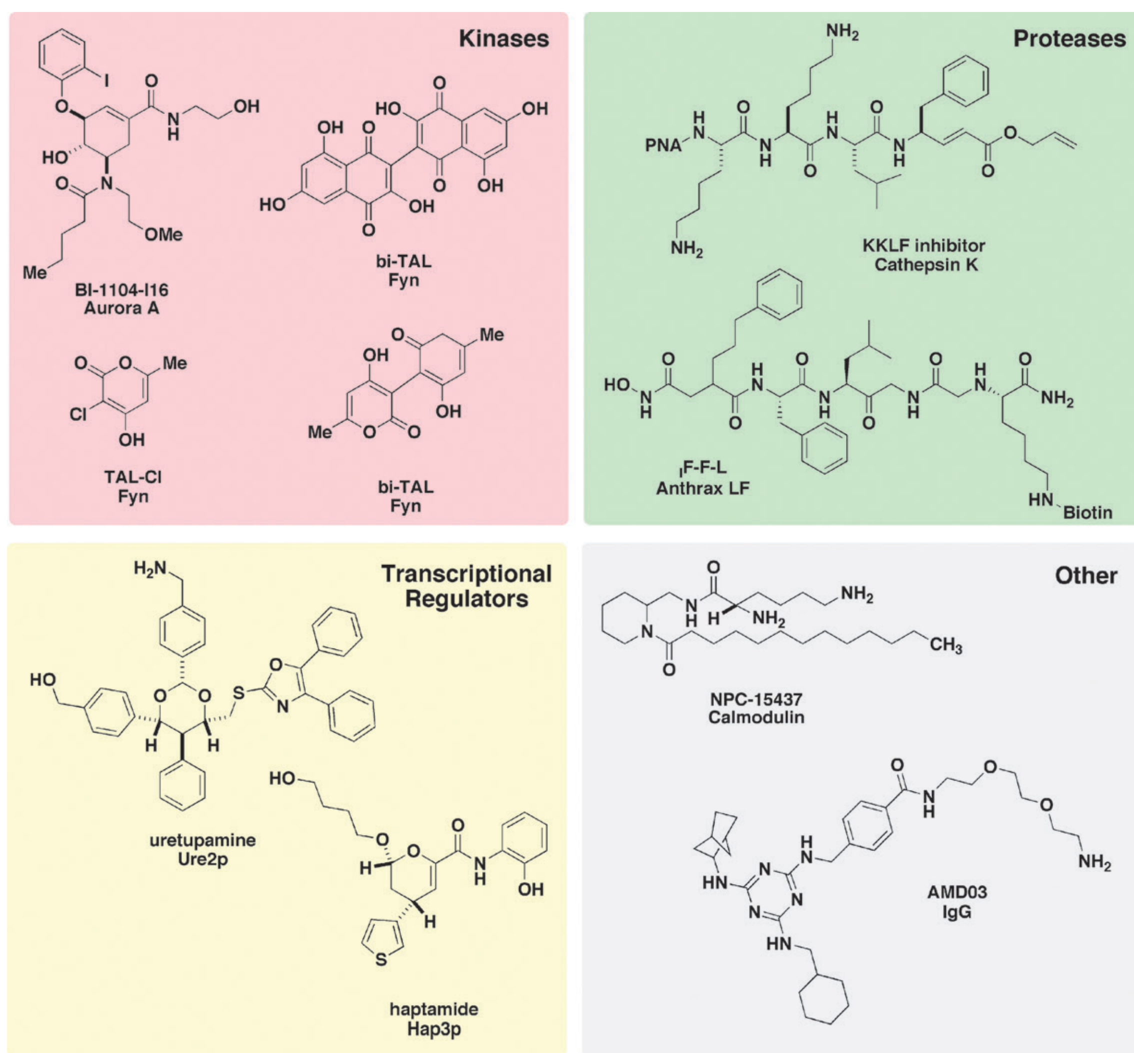


Fig. 4 Representative protein–small molecule interactions discovered using SMMs.[12,21–25,33,45]

preferentially when Ca^{2+} is present in the incubation buffer.[12] Selective inhibitors of closely related cysteine proteases, cathepsin F and cathepsin K, were identified using the PNA-encoded tetrapeptide acrylate microarrays prepared by Winssinger and co-workers.[23] Yao and co-workers screened several metalloproteases against SMMs containing a synthetic hydroxamate peptide library to generate binding signatures for comparisons and to successfully identify lead compounds with affinities in the low micromolar range.[45] Small molecule ligands have also been identified for human IgG,[33] FKBP12,[46] and TNF-alpha.[47] Interactions between RNA secondary structure motifs and small molecules have also been studied using SMMs.[29] Examples of SMM ligand discovery efforts aimed histone deacetylases will be reviewed in more detail to illustrate key concepts relating to SMM preparation, screening, and data analysis.

Histone deacetylases (HDACs) are considered to be valuable therapeutic targets due to their fundamental role in transcriptional regulation and implication in several diseases.[48,49] HDACs catalyze the hydrolysis of *N*-acetyl groups on lysine residues found in the *N*-terminal tails of histone proteins. Small molecules that bind to the various HDACs, especially in a specific manner, may provide additional information about the cellular roles of the various enzymes and serve as templates for therapeutic design.

Several small molecule probes of HDAC function exist and have been reviewed elsewhere.[48] Naturally occurring inhibitors contain common structural features that have guided design of synthetic analogues. The majority of these compounds contain three structural elements that are congruent with proposed structural models for inhibitor binding: a cap region, a linker region, and a metal chelator group.[50,51] The metal chelator element interacts with a catalytic zinc atom while the linker and cap region interact with residues lining the tubular pocket and the opening to the active site, respectively. Suberoylanilide hydroxamic acid (SAHA), recently approved by the FDA for the treatment of cutaneous T-cell lymphoma, is an example of a synthetic compound that follows this structural model and inhibits multiple members of the HDAC family of enzymes.[52] While guidelines are emerging that aid the design of HDAC inhibitors, significant gaps remain in our understanding of which structural features and functionalities allow chemists to design the most potent and selective inhibitors. HTS approaches may be useful in filling out structure–activity

 This journal is © The Royal Society of Chemistry 2008

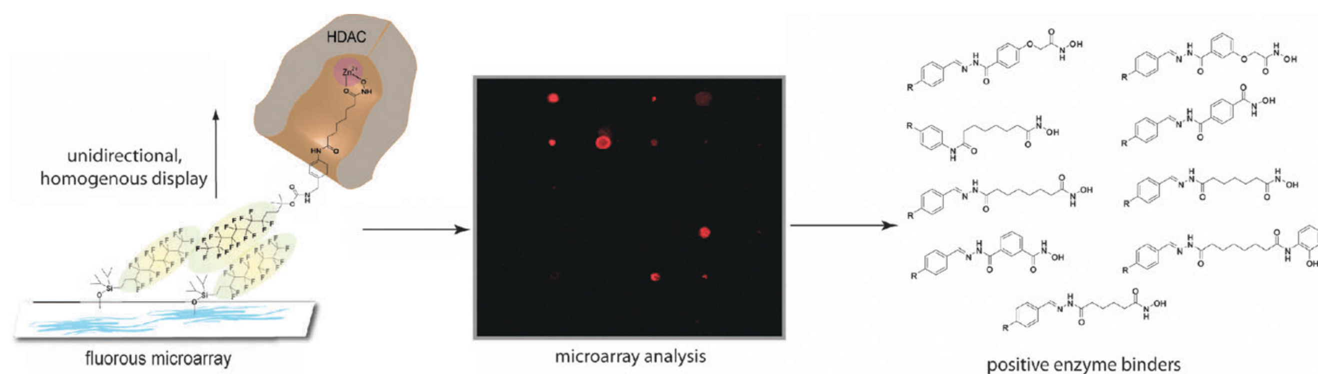


Fig. 5 Discovering inhibitors using HDAC-biased SMMs. Fluorous-tagged small molecules were printed onto fluorous-coated glass slides so that metal chelator groups could be displayed properly for recognition by the enzymes (left). Printed SMMs were incubated with individual HDACs. Putative binders identified by fluorescence intensity analysis (center) were evaluated in enzymatic inhibition assays. The structures for HDAC inhibitors identified using SMM are shown (right).

relationships surrounding existing scaffolds and in identifying new molecular scaffolds that are class specific or specific to individual HDACs.

Schreiber and co-workers screened various HDACs against SMMs containing compounds biased toward HDAC inhibition in an effort to elucidate structure–activity relationships.[18] The authors adopted the non-covalent fluorous SMM approach as it allowed a way to synthesize HDAC-biased libraries in solution and print the compounds onto microarrays with uniform display of a metal chelator group. In this case, orientation is important because the chelator element must reach within the enzyme (Fig. 5). Additionally, uniform orientation provides more confidence in on-array structure–binding relationships. The arrays contained fluorous-tagged compounds with varied linkers and metal chelators as well as negative and positive controls, including fluorous-tagged SAHA.

The arrays were printed using a microcontact robotic arrayer and incubated with purified 6xHis-tag fusions of HDAC2, HDAC3–NCoR2 peptide complex, or HDAC8. SMM assays were run using buffer conditions common to HDAC enzyme activity assays. Interactions between the HDACs and printed compounds were detected using an Alexa-647-labeled antibody against the epitope tag. Quantitative fluorescence data were acquired using a standard fluorescent slide scanner and used to generate a list of putative binders. The authors also performed a competitive on-array binding assay with HDAC3–NCoR2 by including free SAHA in the assay buffer. This condition led to a significant reduction in binding of the complex to the SAHA analogues. Untagged equivalents of the compounds were then tested in enzymatic activity assays with the same set of enzymes to evaluate inhibition. Thermodynamic and kinetic binding data were also collected for untagged compounds binding to one of the HDACs using surface plasmon resonance (SPR) methods. This approach led to the identification of nine compounds that inhibited one or more of the three HDACs with submicromolar IC_{50}'s (Fig. 5).

Finally, the HDAC-biased fluorous SMMs were used in binding assays involving HDACs residing in within whole-cell lysates.[13,18] Most HDACs are thought to reside in multi-protein complexes and require interaction with other proteins for optimal enzymatic activity.[48] SMMs were probed with lysates from 293-MSR cells followed by mouse monoclonal anti-HDAC3 antibody mixed with Alexa-647 labeled secondary antibody. Six of the seven positives on the lysate-based assays also classified as positives with purified HDAC3–NCoR2. Lysate-based SMM assays may provide a more desirable route to identifying small molecule probes of HDAC function as the proteins are screened in a more biologically relevant state. This assay format may be limited by the requirement for specific antibodies of high quality for detecting the protein of interest.

Much effort is placed on structure-aided design of HDAC inhibitors invoking the three-part structural-component model for active-site binding.[18,48–52] Despite this effort, the majority of inhibitors are not isoform selective.[53] Rational design of isoform-selective inhibitors has proven difficult due to the high degree of similarity between active sites among the various enzymes and the lack of structural information about most of the proteins.[48] Novel and unbiased molecular scaffolds may prove useful in selectively inhibiting individual HDACs. With this goal in mind, our group recently initiated a project involving the use of SMMs to screen class I and class II HDACs for novel ligands. Nearly 20 000 diverse compounds were printed on SMMs using isocyanate-mediated capture as describe previously.[12,13] The enzymes used for screening were commercially available as epitope-tagged fusions involving either a 6xHis tag or a glutathione *S*-transferase (GST) tag. The tags allowed fluorescence-based detection of binding using either Alexa 647-labeled antibodies against the epitope tags. Selected HDACs were screened both in the presence or absence of free SAHA.[18] SAHA was added to the incubation buffer in an effort to distinguish SMM positives that are expected to bind at the enzyme active site from potential allosteric binders. In future assays, we hope to screen the same SMMs against lysates that contain various HDACs with appropriate antibodies. Fluorescence images were collected for each replicate and signal-to-noise ratios were computed for each array feature using standard microarray scanner software (Fig. 6a). *Z*-scores, otherwise known as standard scores, were computed for each array feature using a standard analysis

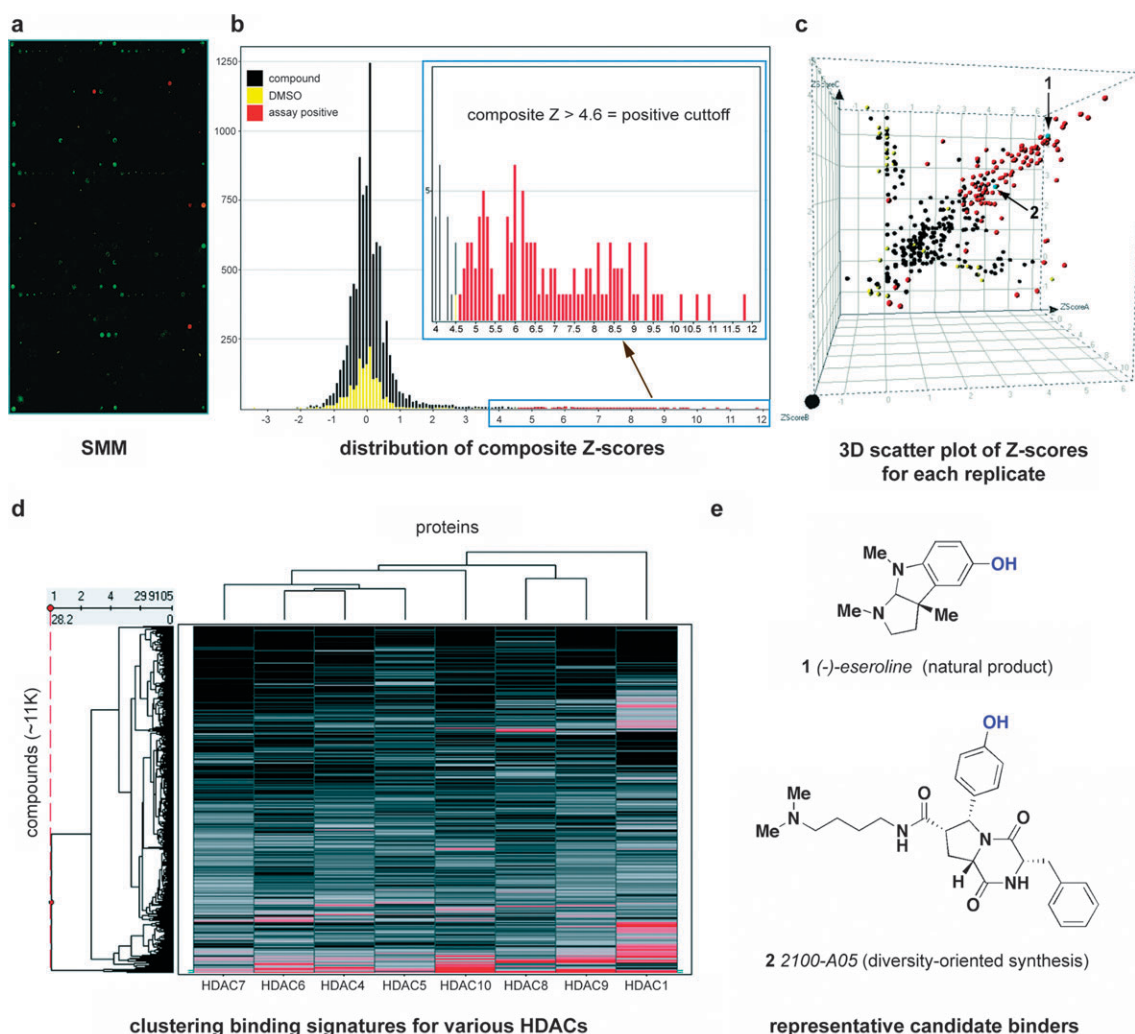


Fig. 6 Unbiased screens for HDAC ligands using SMMs. (a) Image of scanned SMM incubated with purified HDAC8 followed by an Alexa-647 labeled anti-His antibody. Putative HDAC8–small molecule interactions are false-colored red. Alexa-532 control dyes and compounds that autofluoresce at 532 nm are false-colored green. (b) Distribution of composite Z-scores (x-axis) for nearly 11 000 printed features screened against HDAC8 in triplicate. Printed DMSO controls are shown in yellow while compounds are shown in black. Features with composite Z-scores greater than 4.6, shown in red, were judged to be positive in the assay. (c) Three dimensional scatter plot of individual Z-scores for the three replicate assays. This plot is useful for visually evaluating reproducibility. Assay positives are colored red. (d) Binding signatures for various purified HDACs against a common set of nearly 11 000 printed features are shown in a heat map. Compounds with higher composite Z-scores (>4) appear red (composite Z-score thresholds: black = −1, grey = 1, red = 4). Assay positives are evaluated for specificity across the panel of proteins. (e) Two representative positives for HDAC8 (also shown in (c)). Sites of immobilization to the slide are colored blue. The compounds are candidates for additional follow up studies including secondary binding assays, enzymatic assays, or cell-based phenotypic assays.

pipeline for HTS and SMM data developed by Clemons and co-workers.[54] The Z-score indicates how many standard deviations an observation is above or below the mean and allows comparison of observations from different normal distributions. As each screen involved three or more replicates, these replicates were combined to produce a composite Z-score for each compound.[12,54] For each protein, the composite Z-score distribution is viewed as a histogram and thresholds are set to judge SMM assay positives (Fig. 6b). Additionally, scatter plot views of individual Z-scores are useful in evaluating the reproducibility of any given positive (Fig. 6c). Reproducible positives will fall along a line of reproducibility between all three Z-scores rather than skew toward the axis of one particular replicate. After several proteins have been screened against common SMMs, heat maps representing composite Z-scores for a series of compounds and proteins are generated. A partial heat map for the HDAC SMM screen is shown (Fig. 6d). Heat map analysis is useful in judging specificity and can be used to perform structure–binding analyses. Structures for representative positives to HDAC8, including compounds from both synthetic and natural sources, are shown (Fig. 6e). Selected compounds of interest may be evaluated in additional assays including enzymatic assays, SPR assays, and phenotypic assays.[12] It's important to note that specificity in a binding assay is not necessarily predictive of specificity in a cellular context. The SMM binding signatures for the HDACs should be viewed more as a relatively rapid approach to identifying novel binders with annotation regarding *in vitro*

 This journal is © The Royal Society of Chemistry 2008

specificity. Non-selective SMM positives may still exhibit cellular selectivity while isoform-selective HDAC inhibitors may have additional off-target effects in a cell. Currently, we are evaluating selected SMM positives of interest to each of the HDACs in enzymatic and phenotypic assays. Once completed, composite *Z*-score data for the full HDAC SMM screen will be made publicly available on *ChemBank* (http://chembank.broad.harvard.edu/), a public web-based informatics environment developed at the Broad Institute to assist in the analysis of screening data and cheminformatics. In this fashion, we hope to build a chemical biology resource for the HDAC research community.

4. Concluding remarks

Small-molecule microarrays (SMMs) have proven to be a robust and general tool for ligand discovery. Many laboratories have developed novel methods of SMM manufacturing or new screening approaches. SMM assays should prove more useful in the years to come as a wave of new protein targets arising from gene association studies, RNAi studies, phenotypic screening and systematic target identification studies are on the horizon. Many of these targets will not have known structures or functions and will not be easily subjected to functional HTS screens without significant assay development. SMMs can provide a general approach to screening nearly any soluble protein in the soluble proteome. Key technological advances for SMMs should include developing standard operating conditions for screens involving membrane-bound proteins and on-array detection using mass spectrometry. Ligands to targets of interest identified using SMMs or new general binding assays yet to be developed may serve as tool compounds for chemical biology studies or serve as leads in therapeutic development.

Acknowledgements

The authors wish to thank members of the Broad Institute's Chemical Biology Platform for technical and informatics support. Work described herein has been funded in part with Federal funds from the National Cancer Institute's Initiative for Chemical Genetics (Contract No. N01-CO-12400). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Service, nor does the mention of trade names, commercial products or organizations imply endorsement by the US government.

References

1. M. Clamp, B. Fry, M. Kamal, X. Xie, J. Cuff, M. F. Lin, M. Kellis, K. Lindblad-Toh and E. S. Lander, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 19428–19433.
2. G. Wang and T. A. Cooper, *Nat. Rev. Genet.*, 2007, **8**, 749–761.
3. E. N. Ferreira, P. A. F. Galante, D. M. Carraro and S. J. de Souza, *Mol. Biosyst.*, 2007, **3**, 473–477.
4. G. Walsh and R. Jefferis, *Nat. Biotechnol.*, 2006, **24**, 1241–1252.
5. A. Bairoch and R. Apweiler, *Nucleic Acids Res.*, 2000, **28**, 45–48.
6. I. Tirosh, Y. Bilu and N. Barkai, *Curr. Opin. Biotechnol.*, 2007, **18**, 371–377.
7. D. C. Dominguez, R. Lopes and M. L. Torres, *Clin. Lab. Sci.*, 2007, **20**, 234–238.
8. S. L. Schreiber, *Chem. Eng. News*, 2003, **81**, 51–61.
9. D. R. Spring, *Chem. Soc. Rev.*, 2005, **34**, 472–482.
10. R. L. Nicholson, M. Welch, M. Ladlow and D. R. Spring, *ACS Chem. Biol.*, 2007, **2**, 24–30.
11. G. C. Terstappen, C. Schlüpen, R. Raggiaschi and G. Gaviraghi, *Nat. Rev. Drug Discovery*, 2007, **6**, 891–903.
12. J. L. Duffner, P. A. Clemons and A. N. Koehler, *Curr. Opin. Chem. Biol.*, 2007, **11**, 74–82.
13. J. E. Bradner, O. M. McPherson and A. N. Koehler, *Nature Protocols*, 2006, **1**, 2344–2352.
14. G. MacBeath, A. N. Koehler and S. L. Schreiber, *J. Am. Chem. Soc.*, 1999, **121**, 7967–7968.
15. N. Winssinger, Z. Pianowski and F. Debaene, *Top. Curr. Chem.*, 2007, **278**, 311–349.
16. T. Horlacher and P. H. Seeberger, *OMICS*, 2006, **10**, 490–498.
17. K. Schmitz, S. J. Haggarty, O. M. McPherson, J. Clardy and A. N. Koehler, *J. Am. Chem. Soc.*, 2007, **129**, 11346–11347.
18. A. J. Vegas, J. E. Bradner, W. Tang, O. M. McPherson, E. F. Greenberg, A. N. Koehler and S. L. Schreiber, *Angew. Chem., Int. Ed.*, 2007, **46**, 7960–7964.
19. J. E. Bradner, O. M. McPherson, R. Mazitschek, D. Barnes-Seeman, J. P. Shen, J. Dhaliwal, K. E. Stevenson, J. L. Duffner, S. B. Park, D. S. Neuberg, P. T. Nghiem, S. L. Schreiber and A. N. Koehler, *Chem. Biol.*, 2006, **13**, 493–504.
20. N. Kanoh, M. Kyo, K. Inamori, A. Ando, A. Asami, A. Nakao and H. Osada, *Anal. Chem.*, 2006, **78**, 2226–2230.
21. H. Miao, J. A. Tallarico, H. Hayakawa, K. Münger, J. L. Duffner, A. N. Koehler, S. L. Schreiber and T. A. Lewis, *J. Comb. Chem.*, 2007, **9**, 245–253.
22. S. J. Kwon, M. Y. Lee, B. Ku, D. H. Sherman and J. S. Dordick, *ACS Chem. Biol.*, 2007, **2**, 419–425.
23. H. D. Urbina, F. Debaene, B. Jost, C. Bole-Feysot, D. E. Mason, P. Kuzmic, J. L. Harris and N. Winssinger, *ChemBioChem*, 2006, **7**, 1790–1797.
24. F. G. Kuruvilla, A. F. Shamji, S. M. Sternson, P. J. Hergenrother and S. L. Schreiber, *Nature*, 2002, **416**, 653–657.
25. A. N. Koehler, A. F. Shamji and S. L. Schreiber, *J. Am. Chem. Soc.*, 2003, **125**, 8420–8421.
26. P. J. Hergenrother, K. M. Depew and S. L. Schreiber, *J. Am. Chem. Soc.*, 2000, **122**, 7849–7850.
27. J. R. Falsey, M. Renil, S. Park, S. Li and K. S. Lam, *Bioconjugate Chem.*, 2001, **12**, 346–353.
28. F. Fazio, M. C. Bryan, O. Blixt, J. C. Paulson and C.-H. Wong, *J. Am. Chem. Soc.*, 2002, **124**, 14397–14402.
29. J. L. Childs-Disney, M. Wu, A. Pushechnikov, O. Aminova and M. D. Disney, *ACS Chem. Biol.*, 2007, **2**, 745–754.
30. B. T. Houseman and M. Mrksich, *Chem. Biol.*, 2002, **9**, 443–454.
31. D. Barnes-Seeman, S. B. Park, A. N. Koehler and S. L. Schreiber, *Angew. Chem., Int. Ed.*, 2003, **42**, 2376–2379.
32. M. Köhn, R. Wacker, C. Peters, H. Schröder, L. Soulère, R. Breinbauer, C. M. Niemeyer and H. Waldmann, *Angew. Chem., Int. Ed.*, 2003, **42**, 5830–5834.
33. M. Uttamchandani, D. P. Walsh, S. M. Kheronsky, X. Huang, S. Q. Yao and Y. T. Chang, *J. Comb. Chem.*, 2004, **6**, 862–868.
34. M. R. Lee and I. Shin, *Angew. Chem., Int. Ed.*, 2005, **44**, 2881–2884.
35. S. Park, M. R. Lee and I. Shin, *Nature Protocols*, 2007, **2**, 2747–2758.
36. S. Li, N. Marthandan, D. Bowerman, H. R. Garner and T. Kodadek, *Chem. Commun.*, 2005, 581–583.
37. N. Kanoh, A. Asami, M. Kawatani, K. Honda, S. Kumashiro, H. Takayama, S. Simizu, T. Amemiya, Y. Kondoh, S. Hatakeyama, K. Tsuganezawa, R. Utata, A. Tanaka, S. Yokoyama, H. Tashiro and H. Osada, *Chem.–Asian J.*, 2006, **1**, 789–797.
38. Z. Pei, H. Yu, M. Theurer, A. Walden, P. Nilsson, M. Yan and O. Ramström, *ChemBioChem*, 2007, **8**, 166–168.
39. K. S. Ko, F. A. Jaipuri and N. L. Pohl, *J. Am. Chem. Soc.*, 2005, **127**, 13162.
40. R. L. Nicholson, M. L. Ladlow and D. R. Spring, *Chem. Commun.*, 2007, 3906–3908.
41. M. Kurosu and W. A. Mowers, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 3392–3395.
42. J. J. Díaz-Mochón, G. Tourniaire and M. Bradley, *Chem. Soc. Rev.*, 2007, **36**, 449–457.

43. H. Zhu, J. F. Klemic, S. Chang, P. Bertone, A. Casamayor, K. G. Klemic, D. Smith, M. Gerstein, M. A. Reed and M. Snyder, *Nat. Genet.*, 2000, **26**, 283–289.
44. K. Y. Horiuchi, Y. Wang, S. L. Diamond and H. Ma, *J. Biomol. Screening*, 2006, **11**, 48–56.
45. M. Uttamchandani, W. L. Lee, J. Wang and S. Q. Yao, *J. Am. Chem. Soc.*, 2007, **129**, 13110–13117.
46. P. A. Clemons, A. N. Koehler, B. K. Wagner, T. G. Sprigings, D. R. Spring, R. W. King, S. L. Schreiber and M. A. Foley, *Chem. Biol.*, 2001, **8**, 1183–1195.
47. S. E. Tully, M. Rawat and L. C. Hsieh-Wilson, *J. Am. Chem. Soc.*, 2006, **128**, 7740–7741.
48. J. E. Bolden, M. J. Peart and R. W. Johnstone, *Nat. Rev. Drug Discovery*, 2006, **5**, 769–784.
49. N. Khan, M. Jeffers, S. Kumar, C. Hackett, F. Boldog, N. Khramtsov, X. Qian, E. Mills, S. C. Berghs, N. Carey, P. W. Finn, L. S. Collins, A. Tumber, J. W. Ritchie, P. B. Jensen, H. S. Lichenstein and M. Sehested, *Biochem. J.*, 2008, **409**, 581–589.
50. M. S. Finnin, J. R. Donigian, A. Cohen, V. M Richon, R. A. Rifkind, P. A. Marks, R. Breslow and N. P. Pavletich, *Nature*, 1999, **401**, 188–193.
51. A. Vannini, C. Volpari, P. Gallinari, P. Jones, M. Mattu, A. Carff, R. De Francesco, C. Steinkühler and S. Di Marco, *EMBO Rep.*, 2007, **8**, 879–884.
52. P. A. Marks, *Oncogene*, 2007, **26**, 1351–1356.
53. B. E. Schultz, S. Misialek, J. Wu, J. Tang, M. T. Conn, R. Tahilramani and L. Wong, *Biochemistry*, 2004, **43**, 11083–11091.
54. K. Petri Seiler, G. A. George, M. P. Happ, N. E. Bodycombe, H. A. Carrinski, S. Norton, S. Brudz, J. P. Sullivan, J. Muhlich, M. Serrano, P. Ferraiolo, N. J. Tolliday, S. L. Schreiber and P. A. Clemons, *Nucleic Acids Res.*, 2008, **36**, D351–D359.

This journal is © The Royal Society of Chemistry 2008