J|A|C|S COMMUNICATIONS

Published on Web 10/29/2009

Small Molecule-Based Binding Environments: Combinatorial Construction of Microarrays for Multiplexed Affinity Screening

Rachel L. Weller Roska, Tenzing Gawa Surshar Lama, Jay P. Hennes, and Robert E. Carlson*

RECEPTORS LLC, 1107 Hazeltine Boulevard, Suite 510, Chaska, Minnesota 55318

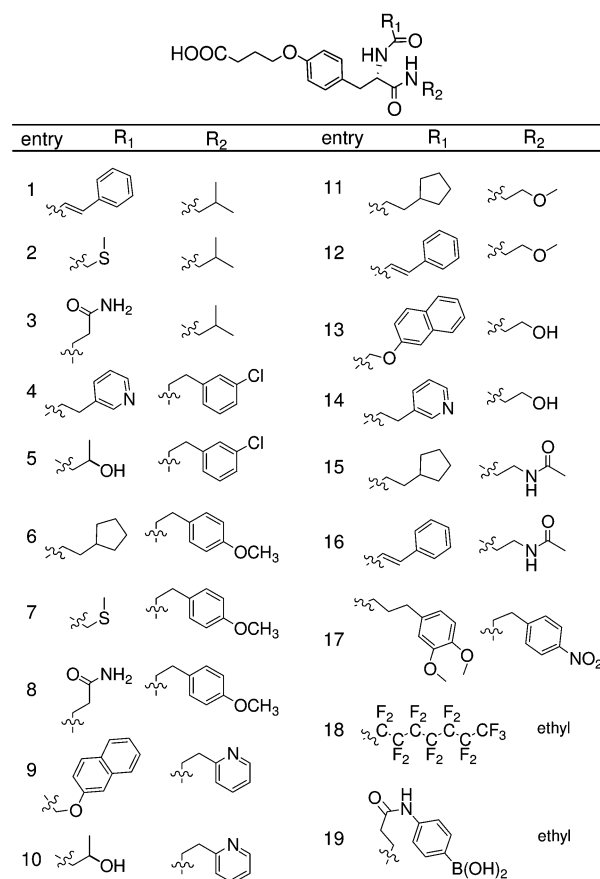
Received June 9, 2009; E-mail: bc@receptorsllc.com

The preparation of stable and diverse binding agents in an addressable format is an objective with myriad applications in proteomics, diagnostics, pharmaceutical development, and related fields. Significant advances toward this aim have been made, and among them microarray-based technologies stand out on the basis of a high-throughput and miniaturized format. In particular, small molecule microarrays[1] (SMMs) have contributed to the expansion of our capabilities in high-throughput screening. The use of synthetic molecules as capture agents bypasses the limitations of approaches that rely on inherently unstable biomolecules, such as antibodies, which can also be laborious and expensive to prepare.[2] The number and variety of small molecules available for immobilization in microarray format have vastly increased through the development of diversity-oriented synthesis.[3] However, the diversity and specificity of the most prominent microarray capture agents, antibodies,[2c] are still unparalleled.[4] Pattern-based molecular recognition, which takes into account the composite binding response of a target across an entire array, has taken us a step closer to resolving this problem and extends the utility of the SMM beyond straightforward lead discovery.[5] However, these microarray strategies still require that each binding agent be discretely synthesized and they principally rely on monovalent interactions with the target.

Microarrays composed of multicomponent,[6] small molecule-based binding *environments* would more closely mimic receptor-like polyvalent binding and facilitate rapid and flexible access to the breadth and depth of binding space. We report here a combinatorial artificial receptor array (CARA) constructed from individual small molecule building blocks immobilized in combinations to form unique and diverse binding environments. A small but diverse library of building blocks (Chart 1, **1**−**19**), designed to incorporate a range of size, polarity, log *P*, and aromaticity, was synthesized. The strategy behind building block design was to mimic and expand upon the functional group diversity of nature's library of amino acids, for example, as presented in an antibody binding pocket. Our library is naive to the target analyte. Each building block was equipped with a pendant carboxylic acid, allowing mixtures of the building blocks to be spotted onto the surface of an amine functionalized slide for covalent immobilization as subunits of the binding environment. In this way, the surface[7] serves as an efficient platform for the combinatorial construction of thousands of discrete binding environments from a concise library of individual small molecules, eliminating the need to discretely synthesize thousands of individual binding agents.

Microarrays displaying 5035 unique binding environments were prepared using our library of 19 small molecule building blocks. The pendant carboxylic acids of the individual building blocks were activated as sulfo-NHS esters, followed by preparation of solutions containing individual activated building blocks or combinations of two, three, or four building blocks, which were then spotted onto the surface of the amine-functionalized glass slide (Figure 1). This array configuration is referred to as an $N_{19}n_{1-4}$ array (19 individual

Chart 1. Building Block Library



building blocks arrayed in combinations displaying one, two, three, or four different building blocks in each binding environment (microarray "spot")). Each binding environment was printed in duplicate, along with control spots for a total of 10 468 spots per microarray. The binding environments are covalently formed and stable,[8] and the building blocks present in every binding environment are known. This strategy contrasts with existing methods[1,5,7] which are dependent on the production of molecular libraries equal in size to the desired number of capture agents for an array, including other methods which rely on *in situ* synthesis on the surface. Moreover, many of these methods require a deconvolution step to identify the compounds printed in a particular spot.[1c]

Four fluorescently labeled proteins, ubiquitin, myoglobin, α-1-acid glycoprotein, and lysozyme, were incubated with $N_{19}n_{1-4}$ arrays to demonstrate the reproducibility of binding and the differentiation of analytes that can be achieved with CARA binding environments. Each array was incubated with 100 nM protein and scanned using a GenePix Personal 4100A fluorescence microarray scanner (Axon Instruments). Figure 2a shows a fluorescence scan of one entire

10.1021/ja9046944 CCC: $40.75
© 2009 American Chemical Society

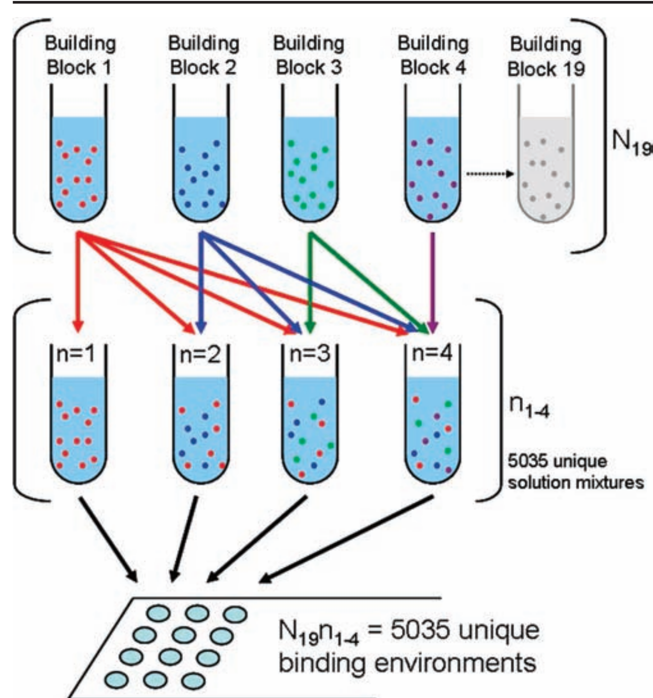


Figure 1. A schematic representation of the CARA building block combination and printing process.

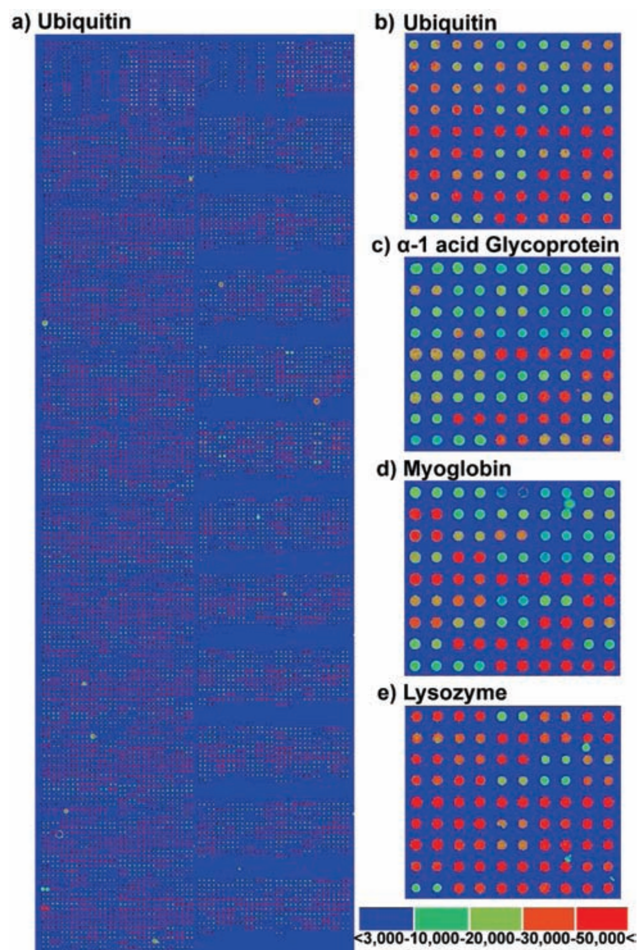


Figure 2. Fluorescence scans of incubated microarrays. (a) Full array incubated with ubiquitin. Zooms of one section of a microarray incubated with (b) ubiquitin, (c) α-1-acid glycoprotein, (d) myoglobin, and (e) lysozyme. Stray flecks of fluorescence visible outside the printed spots in (a), (d), and (e) are artifacts, most likely due to imperfections in the slide surface.

array that was incubated with ubiquitin. The range of binding across the binding environments spans from 1779 to 65 535 fluorescence units (FU). Lower protein concentrations can be used, and a similar range of fluorescence signal can be achieved by adjusting the instrument gain (data not shown). A reduction in binding pattern diversity was typically observed with incubation concentrations below 15 nM (pg/mL range) (data not shown). Figure 2b−e shows one segment of the $N_{19}n_{1-4}$ array from each of four individual protein incubations. The binding differences of the proteins are evident upon simple visual inspection of these fluorescence scans.

Quantitation of the binding patterns for each of the proteins further illustrates the binding differences and highlights the reproducibility of binding, as well as the reproducibility of array construction. Quadruplicate incubations were completed for each protein, and the fluorescence values for each binding environment were quantitated using GenePix Pro 6.1 software. Figure 3a shows a scatter plot comparing the binding data for the first of each duplicate spot on a ubiquitin-incubated array vs the second duplicate spot on the same array. Binding reproducibility is exemplified by the correlation coefficient, 0.99, which is representative of those observed for the other protein incubations (data not shown). A scatter plot illustrating the reproducibility of binding across two separate array incubations for ubiquitin is shown in Figure 3b. Corresponding scatter plots for the other three proteins are shown in Figures S2−S4 in the Supporting Information. The correlation coefficients for each protein were 0.98 for ubiquitin, 0.97 for α-1-acid glycoprotein, 0.98 for myoglobin, and 0.96 for lysozyme. This high correlation of binding across duplicate arrays clearly demonstrates the reproducibility of our approach to binding environment construction. The fluorescence units from the quadruplicate incubations for each protein were averaged,[9] and scatter plots comparing binding data for each protein vs the other three proteins are shown in Figure 3c−h. It is apparent that the binding patterns for the proteins are dramatically different. There is a range of differentiation among the protein pairings, with some of the binding patterns differing more than others. In the case of two more similar binding patterns, such as those of ubiquitin and myoglobin (Figure 3h), there are still numerous binding environments within the array that distinguish the proteins. These data clearly demonstrate the capacity of the combinatorial binding environments to differentiate proteins via their binding patterns, which underscores the potential value the arrays hold for a variety of applications.

An advantage of the CARA strategy is the enormous flexibility it enables in the preparation of alternate microarray configurations, which can be tailored for specific applications. A recognized bottleneck in microarray technology is synthesis and display of binding agents.[1b] Addition of a single building block to the $N_{19}n_{1-4}$ array configuration would result in an $N_{20}n_{1-4}$ array affording 1159 new binding environments, for a total of 6195. For some applications it may be advantageous to probe binding environments of greater complexity (more building blocks per binding environment). Arraying the original 19 building blocks in combinations of up to five building blocks per environment (an $N_{19}n_{1-5}$ array) would yield 16 663 binding environments. Alternatively, binding environments with increased chemical complexity can be screened through the sequential use of two array configurations. An $N_{29}n_{1-2}$ array (435 binding environments) displays 29 building blocks singly or in pairs. This array allows a wider diversity of chemical functionality to be screened through the use of a larger building block library, while maintaining a manageable number of binding environments. Binding data from this array can be used to select the building blocks that

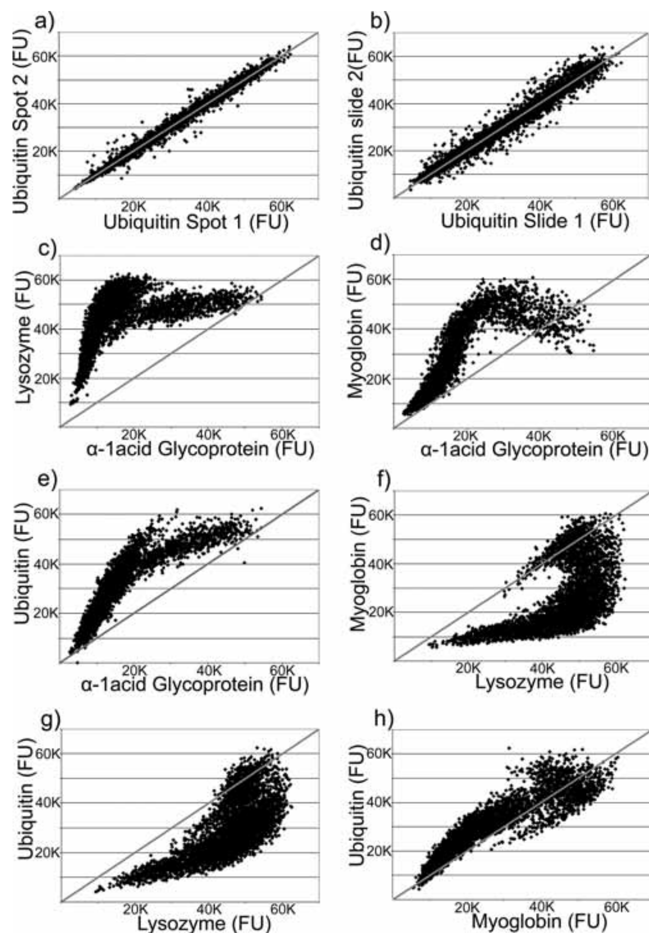


Figure 3. Scatter plots comparing binding data for (a) duplicate spots from one microarray incubated with ubiquitin, (b) two separate ubiquitin incubations, (c) α-1-acid glycoprotein vs lysozyme incubations, (d) α-1-acid glycoprotein vs myoglobin incubations, (e) α-1-acid glycoprotein vs ubiquitin incubations, (f) lysozyme vs myoglobin incubations, (g) lysozyme vs ubiquitin incubations, and (h) myoglobin vs ubiquitin incubations.

exhibit the desired affinity characteristics. These building blocks are then arrayed in a more focused, but higher order configuration, such as an N_9n_{1-9} array (511 binding environments). Using arrays in tandem facilitates access to a depth of binding space that is comparable to, but more targeted than, a single array that incorporates more building blocks in more combinations. For example, an $N_{29}n_{1-9}$ array would require printing and analysis of greater than 16 million binding environments. This sequential array workflow also provides the advantage of preselection of targeted binding characteristics using the first array and further fine-tuning of those binding results in the second array.

Ongoing work in our lab is focused on application of CARA microarrays to a wider variety of analytes and adaptation of binding environments to supports other than glass slides. We are using the sequential array workflow described above to identify lead binding environments for construction of selective affinity purification supports on polymeric and controlled-pore glass beads. We are also using the microarrays to screen for candidate binding environments with binding characteristics appropriate for biosensor applications and for construction on self-assembled monolayers for surface plasmon resonance, as well as surfaces appropriate for mass spectrometric analysis. Nonprotein analytes currently under investigation with CARA microarrays include small molecules, microbes, nucleic acids and cells. The fact that the binding environments are constructed from synthetic small molecules makes them readily scalable to large batch sizes for a wide range of formats and applications.

We constructed combinatorial artificial receptor arrays, which offer significant advantages over existing small molecule microarray strategies, and demonstrated their capacity for protein differentiation. The CARA strategy employs the microarray surface as the combinatorial synthesis platform, which allows for flexibility in array preparation and agility in application. Thousands of unique and diverse binding environments were generated from 19 discretely synthesized building blocks. Binding is reproducible, indicating that array construction is also reproducible, and the diversity of binding across the array shows that a wide range of binding interactions are possible.

Acknowledgment. We are grateful to Robert M. Carlson and Thomas Kodadek for helpful discussions; Robert Kaufman, Bryan Jones, Erika Walden, and Diane Isabell for their contributions to this manuscript; and the University of Minnesota NMR lab.

Supporting Information Available: General synthetic scheme, characterization data for **1**−**19**, experimental procedures for building block activation and printing and protein labeling and incubation. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Macbeath, G.; Koehler, A. N.; Schreiber, S. L. *J. Am. Chem. Soc.* **1999**, *121*, 7967–7968. (b) Uttamchandani, M.; Walsh, D. P.; Yao, S. Q.; Chang, Y. *Curr. Opin. Chem. Biol.* **2005**, *9*, 4–13. (c) Uttamchandani, M.; Wang, J.; Yao, S. Q. *Mol. BioSyst.* **2006**, *2*, 58–68.
(2) (a) Reference 1c. (b) MacBeath, G.; Schreiber, S. L. *Science* **2000**, *289*, 1760–1763. (c) Tomizaki, K.; Usui, K.; Mihara, H. *ChemBioChem* **2005**, *6*, 782–799.
(3) (a) Schreiber, S. L. *Science* **2000**, *287*, 1964–1969. (b) Tan, D. S. *Nat. Chem. Biol.* **2005**, 74–84. (c) Morton, D.; Leach, S.; Cordier, C.; Warriner, S.; Nelson, A. *Angew. Chem., Int. Ed.* **2008**, *48*, 104–109. (d) Hergenrother, P. J.; Depew, K. M.; Schreiber, S. L. *J. Am. Chem. Soc.* **2000**, *122*, 7849–7850. (e) Arya, P.; Joseph, R.; Gan, Z.; Rakic, B. *Chem. Biol.* **2005**, *12*, 163–180.
(4) Collins, B. E.; Wright, A. T.; Anslyn, E. V. *Top. Curr. Chem.* **2007**, *277*, 181–218.
(5) (a) Reference 4. (b) Wright, A. T.; Griffin, M. J.; Zhong, Z.; McCleskey, S. M.; Anslyn, E. V.; McDevitt, J. T. *Angew. Chem., Int. Ed.* **2005**, *44*, 6375–6378. (c) Baldini, L.; Wilson, A. J.; Hong, J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2004**, *126*, 5656–5657. (d) Reddy, M. M.; Kodadek, T. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12672–12677.
(6) (a) Melkko, S.; Scheuermann, j.; Dumelin, C. E.; Neri, D. *Nat. Biotechnol.* **2004**, *22*, 568–574. (b) Miyazaki, I.; Simizu, S.; Ishida, k.; Osada, H. *ChemBioChem* **2009**, *10*, 838–843.
(7) (a) Fodor, S. P.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. *Science* **1991**, *251*, 767–773. (b) Li, S.; Bowerman, D. J.; Marthandan, N.; Garner, H. R.; Kodadek, T. *J. Am. Chem. Soc.* **2004**, *126*, 4088–4089.
(8) The binding characteristics of the arrays did not deteriorate after heating at 80 °C in an oven or boiling in 20% aqueous ethanol for 2 h (data not shown).
(9) The average coefficients of variation (CV) for the binding environments within a quadruplicate set were 8% for ubiquitin, 13% for α-1-acid glycoprotein, 14% for myoglobin, and 9% for lysozyme.

JA9046944