Protein Analysis

DOI: 10.1002/ange.200501137

Differential Receptors Create Patterns That Distinguish Various Proteins**

Aaron T. Wright, Michael J. Griffin, Zhenlin Zhong, Shawn C. McCleskey, Eric V. Anslyn, and John T. McDevitt**

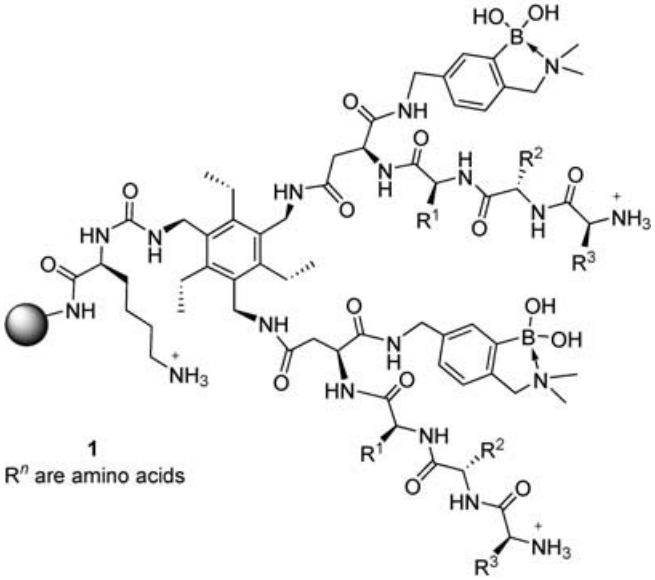
The term "protein-detecting array" has been coined to describe an analytical device consisting of a series of protein receptors.[1] Classically, such an array consists of highly selective protein-binding agents. Unfortunately, the creation of synthetic receptors with high affinity and specificity for proteins is particularly challenging owing to their molecular complexity. However, there has been recent success in creating selective synthetic receptors for assorted peptides.[2]

An alternative to the use of highly selective receptors is to employ differential receptors.[3] This method uses an array of receptors having good affinity, but not necessarily high selectivity, for a particular target. When functioning in tandem the combined response of the receptors creates a pattern that is diagnostic for each analyte. Because the response of the receptor array does not necessarily rely on specific recognition interactions between substrate and analyte, highly challenging analytes can be targeted.

Very few groups have used differential receptors for the analysis of biomolecules.[3] Hamilton et al. recently reported the use of a 35-member receptor library of tetraphenyl porphyrin derivatives, functionalized with different amino acids resulting in charges varying from +8 to −8, for the detection of charged proteins. Eight members of the library were used to detect four charged proteins, ranging from acidic ferredoxin (pI = 2.75) to highly basic cytochrome C (pI = 10.6). Fluorescence quenching patterns composed from the responses of the eight receptors correlated with the charge complementarity between the receptors and proteins.[4] Until our current report, differential receptor arrays had not found use in distinguishing classes of proteins.

Our group's approach to differential sensing is to create libraries of receptors that are biased towards particular analyte classes. Using this approach, we have shown that a microchip-based array incorporating a combinatorial library of receptors[5] is effective in differentiating nucleotide phosphates with an indicator-displacement assay.[6] We now report the development of a library of differential receptors biased towards proteins and glycoproteins. When combined with an indicator-uptake visualization assay and principal component analysis (PCA), the library gives differentiation of proteins and glycoproteins, as well as subtle differentiation within each protein class.[7]

To differentiate proteins and glycoproteins, we designed and synthesized library **1**, which incorporates one of 19



natural amino acids (cysteine excluded) at each of three sites on two different binding arms, by using combinatorial chemistry.[8] This created a library with 19^3 (6859) unique members. The peptide arms provide sites for molecular recognition of proteins by means of ion pairing, hydrogen bonding, and the hydrophobic effect. The boronic acids provide effective sugar binding sites because these groups rapidly and reversibly form cyclic esters with diols in aqueous media.[9] The hexasubstituted benzene scaffold acts as a spacer and assists in the creation of a binding cavity.[10] Our expectation was that each receptor would show differential binding with proteins based on the variance in the peptide arms, and that the boronic acids would assist in differentiation of proteins from glycoproteins.

We used a 7 × 5 array consisting of 29 randomly selected resin beads from library **1** and six acylated resin blanks. Each bead was placed in a micromachined chip-based array platform (see the Supporting Information) that has been previously described.[11] Protocols for the delivery of protein and indicator, and for acidic and basic washes? have been detailed previously. The patterns created by the array of receptors are obtained by measuring the intensity of transmitted red, green, and blue light for each bead using a charge-coupled device attached to a customized reader.[12]

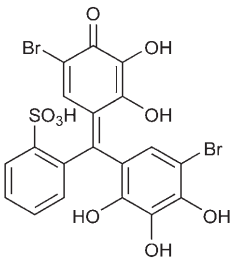
In order to create diagnostic patterns for each protein, a signaling protocol was incorporated. Previous work utilized

[*] A. T. Wright, M. J. Griffin, Dr. Z. Zhong, S. C. McCleskey, Dr. E. V. Anslyn, Dr. J. T. McDevitt
Department of Chemistry and Biochemistry
The University of Texas at Austin
Austin, TX 78712 (USA)
Fax: (+1) 512-471-8696
E-mail: Anslyn@ccwf.cc.utexas.edu
Mcdevitt@mail.utexas.edu

[**] We gratefully acknowledge support for this work from the U.S. NIH (EB00549-5). We thank Dr. Lara Mahal for her advice and insight. Thanks also to Dr. Klaus Linse and Michelle Gadush for their assistance in sequencing beads.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

indicator-displacement assays; however, at the low protein concentrations used in this study it was more advantageous to evaluate the rate of indicator uptake. Our indicator-uptake signaling protocol allowed us to use protein concentrations (355 μM) nearly 60 times less than that in our previous work with nucleotide phosphates (20 mM).[6] We incorporated the commercially available indicator bromopyrogallol red (**2**) for the indicator-uptake colorimetric analysis. This indicator forms reversible cyclic esters with the boronic acids of the receptor.



2

Our experimental protocol starts with a delivery of 5 mL of a protein solution (355 μM in HEPES buffer) at a flow rate of 0.25 mL min^{-1}. This is followed by washing with buffer for 3 min (25 mM HEPES, pH 7.4, 1.0 mL min^{-1}) and then delivery of **2** (3.0 μM in HEPES buffer, 1.0 mL min^{-1}). After each analysis, the protein and indicator were washed from the array with NaOH (150 mM) and HCl (300 mM) rinses. This allowed for repeated use of the array. A 12-bit image was captured every 2 s during indicator uptake. For each receptor bead in the array the green-channel absorbance ($\lambda = 550$ nm) was plotted versus time and a slope was garnered. Effective absorbance values were obtained by calculating the negative log of the ratio of the green-channel intensity of each bead to the green-channel intensity of a blank bead.[11]

Ovalbumin, fetuin, lysozyme, bovine serum albumin (BSA), and elastin were used for this study. These choices were made to challenge our design principles by grouping proteins of similar properties. The characteristics of the proteins span a variety of molecular weights, glycosidic properties, and isoelectric points (pI). The molecular weights of ovalbumin and fetuin are similar, as are those of elastin and BSA. The pI of ovalbumin, BSA, and fetuin are similar, as are those of lysozyme and elastin (Table 1).

Table 1: Characteristic protein data.

Protein	M_w [kDa]	pI	Glycoprotein?
ovalbumin	44–45	4.6	yes
fetuin	48.4	4.5–4.9	yes
lysozyme	14	9.6–11	no
BSA	66	4.7–5.2	no
elastin	60	9.3–10.2	no

Four trials were performed for each protein.[13] The indicator-uptake slopes were calculated for each receptor bead over the time during which the dye passed through the array (49 to 403 s). An illustration of the differences in indicator uptake from different resin-bound receptors can be seen in Figure 1. For each trial a slope was measured for each bead (blanks not included). Because of the large number of slopes calculated from each trial, the dimensionality of the data set was simplified using PCA.

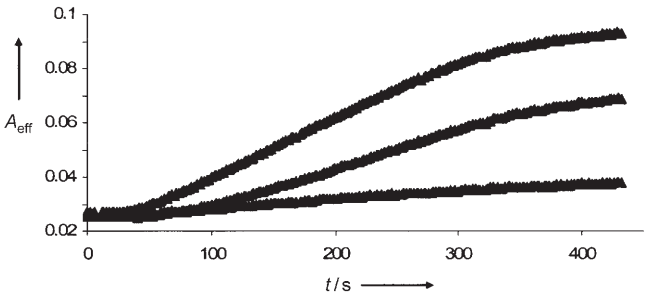


Figure 1. Indicator-uptake curves from three beads are represented from a trial with BSA. The slope from 49 to 403 s for each resin-bound receptor is taken as part of a cumulative slope pattern for each protein. A_{eff} is the effective absorbance. See the text for an explanation of the measurement procedure.

In PCA, the first principal component (PC) axis is calculated to lie along the line of maximum variance in the original data set. Subsequent orthogonal axes are calculated to lie along lines of diminishing levels of variance. In this study, the first four PC axes effectively satisfied the Kaiser criterion, which states that as many factors can be extracted as variables that have eigenvalues greater than one.[14] Figure 2

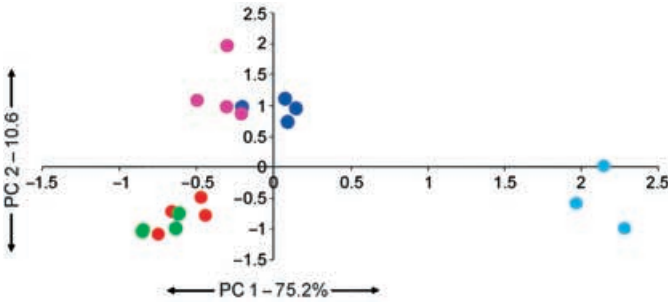


Figure 2. PCA score plot for each protein trial. Percentages on the PC axes define the weight of those axes to the overall pattern. Spacing between proteins demonstrates the recognition capability of the resin-bound receptors. (● lysozyme, ● elastin, ● ovalbumin, ● fetuin, ● BSA)

shows a two-dimensional PCA plot which effectively separates the different protein classes. As illustrated, there is differentiation between proteins and glycoproteins, which was a primary goal of this study. However, because four PC axes are outside the range of error it was possible to generate a three-dimensional PCA plot that further separates the proteins. Figure 3 shows an expanded PCA plot using PC axes 1–3. This demonstrates that the array of receptors adequately separated proteins from glycoproteins, and to a lesser extent even separates proteins within the classes. The fact that we used a random selection of receptors to achieve these results illustrates the power of our method. More than likely any one of these 29 receptors would not have been a good receptor for a particular protein; however, when functioning in tandem, the array works well. Further, any 29 receptors could presumably be used to obtain analogous patterns for the proteins. PC axis 4 (4.7%) could have also been used for further discrimination in another plot.

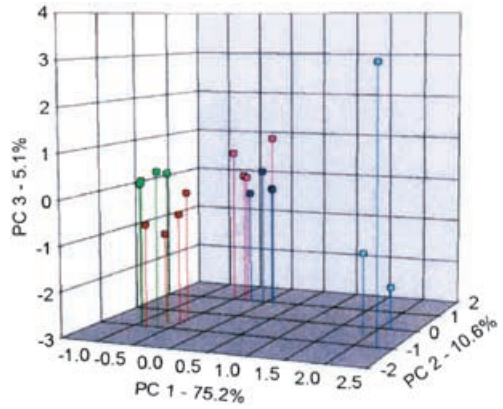


Figure 3. PCA score plot for each protein trial (● lysozyme, ● elastin, ● ovalbumin, ● fetuin, ● BSA). The additional dimension further separates the proteins.

A control was performed using a 7×5 array with six blanks and 29 resin beads derivatized only with tripeptides that were obtained from a combinatorial library synthesized with 19 natural amino acids (cysteine excluded). Using this array of tripeptide beads, no separation of analytes occurred. This control illustrates that some design must be incorporated into the receptors to bind certain analyte classes, and that simple random receptors are inefficient. These results do not strictly demonstrate that the boronic acids bind to glycoproteins and not standard proteins. More likely the boronic acids interact with surface epitopes on both protein classes but to different extents. Yet, if differential formation of covalent bonds with the boronic acids was the only factor in the discrimination between protein classes, then because they are the only binding moiety present in every receptor we would expect no differences within each protein class. Therefore, both the boronic acids and the variable peptide arms of the receptors are critical in the identification and discrimination of proteins and glycoproteins.

Though we significantly reduced the analyte concentrations from those in our previous report with nucleotide phosphates, the protein concentrations (355 μM) are still relatively high for practical analysis. Reduction to practical concentrations (nM) was accomplished with differential receptors of higher affinity. In other work we have also shown that an analysis cell that recycles the analyte solution through the array leads to extremely sensitive assays.[15]

The PCA plot (Figure 2 and Figure 3) demonstrates similarities between ovalbumin and fetuin, and similarities between elastin and lysozyme. Yet, even the proteins in similar groups are separated. The proximity of ovalbumin and fetuin is reasonable as both are glycoproteins with similar pI values, and both likely interacted with the boronic acid moieties. Further, elastin and lysozyme have similar pI values. Therefore, the separation in our analysis is not simply a result of charge differences. Interestingly, molecular weight also did not play a large role in the patterns exhibited. BSA is likely separated from the others because it has a different pI than elastin and lysozyme and is not a glycoprotein. The differentiation between similar proteins is likely due to specific contacts between the receptors and proteins that are cross-reactive and subtly discriminatory.

Factor loading values are calculated in PCA to determine the magnitude of the contribution of an original variable to the formation of a PC axis. Variables with loading values approaching −1 or 1 have a dominant role in the formation of a PC axis. Because PC axis 1 described the most variance, five beads with high loading values on PC 1 and two beads with low loading values were selected for receptor characterization (Table 2) using Edman degradation. The sequencing results do not show any obvious homologies. Yet, the lack of any homology is a lesson in itself: differential sensing schemes can be successful and may even benefit from a wide variety of structurally diverse receptors.

Table 2: Factor loadings and sequencing results for principal component 1.

Tripeptide sequence	Factor loadings (PC 1)	Bead number
Ala-Ser-Asp	0.984	12
Ser-Lys-Gly	0.963	9
Arg-Lys-Lys	0.951	15
Gly-Asp-Ser	0.932	2
Asp-Leu-Val	0.928	22
Lys-Arg-Met	0.774	23
Gly-Gln-Gln	0.722	6

In summary, we have shown that the use of an array of differential receptors can differentiate between classes of proteins, and even between very structurally similar proteins. This separation did not arise from charge differences or molecular weight differences, but rather from specific contacts between receptors and proteins giving discriminatory patterns. We plan to expand this general approach to the qualitative analysis of complex mixtures of proteins.

Received: March 30, 2005
Revised: June 6, 2005
Published online: September 12, 2005

Keywords: host–guest systems · proteins · receptors · supramolecular chemistry

[1] T. Kodadek, *Chem. Biol.* **2001**, *8*, 105–115.
[2] a) A. T. Wright, E. V. Anslyn, *Org. Lett.* **2004**, *6*, 1341; b) A. Buryak, K. Severin, *Angew. Chem.* **2004**, *116*, 4875–4878; *Angew. Chem. Int. Ed.* **2004**, *43*, 4771–4774; c) H. Wennemers, M. Conza, M. Nold, P. Krattiger, *Chem. Eur. J.* **2001**, *7*, 3342–3347; d) M. W. Peczuh, A. D. Hamilton, *Chem. Rev.* **2000**, *100*, 2479–2494; e) W. C. Still, *Acc. Chem. Res.* **1996**, *29*, 155–163.
[3] a) A. Buryak, K. Severin, *J. Am. Chem. Soc.* **2005**, *127*, 3700–3701; b) J. J. Lavigne, E. V. Anslyn, *Angew. Chem.* **2001**, *113*, 3212–3225; *Angew. Chem. Int. Ed.* **2001**, *40*, 3118–3130; c) S. L. Wiskur, P. N. Floriano, E. V. Anslyn, J. T. McDevitt, *Angew. Chem.* **2003**, *115*, 2116–2118; *Angew. Chem. Int. Ed.* **2003**, *42*, 2070–2072; d) K. J. Albert, N. S. Lewis, C. L. Schauer, G. A. Sotzing, S. E. Stitzel, T. P. Vaid, D. R. Walt, *Chem. Rev.* **2000**, *100*, 2595–2626; e) N. A. Rakow, K. S. Suslick, *Nature* **2000**, *406*,

710 – 713; f) T. Hirsch, H. Kettenberger, O. S. Wolfbeis, V. M. Mirsky, *Chem. Commun.* **2003**, 432 – 433.
[4] L. Baldini, A. J. Wilson, J. Hong, A. D. Hamilton, *J. Am. Chem. Soc.* **2004**, *126*, 5656 – 5657.
[5] S. L. Wiskur, H. Ait-Haddou, J. J. Lavigne, E. V. Anslyn, *Acc. Chem. Res.* **2001**, *34*, 963 – 972.
[6] S. C. McCleskey, M. J. Griffin, S. E. Schneider, J. T. McDevitt, E. V. Anslyn, *J. Am. Chem. Soc.* **2003**, *125*, 1114 – 1115.
[7] a) W. Carey, K. Beebe, B. Kowalski, *Anal. Chem.* **1986**, *58*, 149 – 153; b) P. Jurs, G. Bakken, H. McClelland, *Chem. Rev.* **2000**, *100*, 2649 – 2678.
[8] Z. Zhong, E. V. Anslyn, *J. Am. Chem. Soc.* **2002**, *124*, 9014 – 9015.
[9] a) H. G. Kuivila, A. H. Keough, E. J. Soboczenski, *J. Org. Chem.* **1954**, *19*, 780 – 783; b) J. P. Lorand, J. D. Edwards, *J. Org. Chem.* **1959**, *24*, 769 – 774; c) T. D. James, K. R. A. S. Sandanayake, S. Shinkai, *Angew. Chem.* **1996**, *108*, 2051 – 2053; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1910 – 1922; d) S. L. Wiskur, J. J. Lavigne, H. Ait-Haddou, V. Lynch, Y. H. Chiu, J. W. Canary, E. V. Anslyn, *Org. Lett.* **2001**, *3*, 1311 – 1314.
[10] a) D. J. Iverson, G. Hunter, J. F. Blount, J. R. Damewood, K. Mislow, *J. Am. Chem. Soc.* **1981**, *103*, 6073 – 6083; b) K. V. Kilway, J. S. Siegel, *J. Am. Chem. Soc.* **1992**, *114*, 255 – 261; c) H. E. Gottlieb, C. Ben-Ary, A. Hassner, V. Marks, *Tetrahedron* **1999**, *55*, 4003 – 4114.
[11] A. Goodey, J. J. Lavigne, S. Savoy, M. Rodriguez, T. Curey, A. Tsao, G. Simmons, S. Yoo, Y. Sohn, E. V. Anslyn, J. Shear, D. Niekirk, J. T. McDevitt, *J. Am. Chem. Soc.* **2001**, *123*, 2559 – 2570.
[12] a) K. Niikura, A. Metzger, E. V. Anslyn, *J. Am. Chem. Soc.* **1998**, *120*, 8533 – 8534; b) J. J. Lavigne, E. V. Anslyn, *Angew. Chem.* **1999**, *111*, 3903 – 3906; *Angew. Chem. Int. Ed.* **1999**, *38*, 3666 – 3669; c) S. L. Wiskur, E. V. Anslyn, *J. Am. Chem. Soc.* **2001**, *123*, 10109 – 10110; d) See refs. [4] and [5].
[13] One BSA trial was excluded because an air bubble skewed the CCD images.
[14] H. F. Kaiser, *Educ. Psychol. Meas.* **1960**, *20*, 141 – 150.
[15] N. Christodoulides, S. Mohanty, C. S. Miller, M. C. Langub, P. N. Floriano, P. Dharshan, M. F. Ali, B. Bernard, D. Romanovicz, E. Anslyn, P. C. Fox, J. T. McDevitt, *Lab on a Chip* **2005**, *5*, 261 – 269.

