J|A|C|S
ARTICLES
Published on Web 11/03/2006

A Rational Approach to Minimal High-Resolution Cross-Reactive Arrays

Eric Green,[†] Mark J. Olah,[‡] Tatiana Abramova,[†] Lance R. Williams,[‡] Darko Stefanovic,[‡] Tilla Worgall,[†] and Milan N. Stojanovic*,[†]

Contribution from the Division of Clinical Pharmacology and Experimental Therapeutics, Department of Medicine, and Department of Pathology, Columbia University, Box 84, 630 West 168th Street, New York, New York 10032, and Department of Computer Science, University of New Mexico, Albuquerque, New Mexico 87131

Received June 16, 2006; E-mail: mns18@columbia.edu

Abstract: We report a rational approach to the construction of cross-reactive arrays for steroids consisting of five to seven sensors incorporating modified oligonucleotides. The sensors for our arrays were selected to maximize their differential responses to the two steroids most different in an arbitrarily chosen parameter named "shape-length". The arrays incorporated three previously unreported types of sensors identified through a massive screening effort: (1) three-way junction sensors with neutralized charges within junction; (2) "self-aggregating sensors"; and (3) sensors incorporating fluorophore directly in a three-way junction as a spacer. The arrays were tested on seven steroids and an alkaloid (cocaine) over a range of concentrations, and achieved 92−96% accuracy in class assignments, despite the close structural similarities between steroids.

Introduction

Closely related compounds, differing in a single methyl group, or a single stereocenter, can be readily distinguished by a mammalian olfactory system.[1] This is accomplished by an array of about 1000 nonspecific receptors cross-interacting with odorants, and generating patterns characteristic for each odorant.[2] The concept of cross-reactive sensor arrays[3] has been introduced in the hope of reproducing the resolution and sensitivity of the mammalian olfactory system, and has led to important developments, particularly in the analysis of gaseous analytes.[3] In recent years, several impressive examples of cross-reactive arrays of molecular receptors have been reported that target solution-phase analytes as well.[4]

In our previous report in this journal,[5] we demonstrated that the DNA three-way junction[6] provides a scaffold for the systematic generation of cross-reactive sensors capable of signaling the presence of hydrophobic molecules. In that work, we constructed a nine-sensor cross-reactive array by systematically varying the positions of fluorophores and the mismatch content in junctions. However, our approach was purely empirical, many sensors were redundant, and the array did not have sufficient resolution to solve the challenge posed below.

In Figure 1, we show seven steroids representative of various steroid classes, and we pose the following challenge: Given the large number of sensors with known properties, can we develop a rational procedure to find the minimum number of sensors that would form a cross-reactive array capable of generating a fingerprint for each of these steroids over a range of concentrations? While strong arguments have been made for large arrays,[3] we were interested in a small array for several, purely practical reasons. Such arrays would be more cost-effective, they would be easier to train than larger arrays (while avoiding Curse of Dimensionality[7]), they would be sufficient for most common applications, and their results would be easier to transfer to other laboratories.

The task to differentiate steroids in Figure 1 is particularly difficult for artificial molecular-scale receptors with hydrophobic interactions as their primary recognition motifs, because some of these steroids have very similar hydrophobic properties and shapes. Aside from the manifest engineering challenge, any array capable of generating characteristic patterns for these steroids over a wide range of concentrations would also have practical applications in diagnostic tests. In metabolic errors of steroido-

[†] Columbia University.
[‡] University of New Mexico.

(1) Abate, A.; Brenna, E.; Fuganti, C.; Gatti, G. F.; Giovenzana, T.; Malpezzi, L.; Serra, S. *J. Org. Chem.* **2005**, *70*, 1281.
(2) (a) Lledo, P.-M.; Gheusi, G.; Vincent, J.-D. *Physiol. Rev.* **2005**, *85*, 281. (b) Firestein, S. *Nature* **2001**, *413*, 211.
(3) (a) Lewis, N. L. *Acc. Chem. Res.* **2004**, *37*, 663. (b) Albert, K. J.; Lewis, N. S.; Schauer, C. L.; Sotzing, G. A.; Stitzel, S. E. I.; Vaid, T. P.; Walt, D. R. *Chem. Rev.* **2000**, *100*, 2595 and references therein. (c) Lavigne, J. J.; Anslyn, E. V. *Angew. Chem., Int. Ed.* **2001**, *40*, 3118.
(4) (a) Folmer-Andersen, J. F.; Kitamura, M.; Anslyn, E. V. *J. Am. Chem. Soc.* **2006**, *128*, 5652. (b) Zhou, H.; Baldini, L.; Hong, J.; Wilson, A. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2006**, *128*, 2421. (c) Zhang, C.; Suslick, K. S. *J. Am. Chem. Soc.* **2005**, *127*, 11548. (d) Buryak, A.; Severin, K. *J. Am. Chem. Soc.* **2005**, *127*, 3700. (e) Greene, N. T.; Shimizu, K. D. *J. Am. Chem. Soc.* **2005**, *127*, 5695. (f) Goddard, J.-P.; Reymond, J.-L. *J. Am. Chem. Soc.* **2004**, *126*, 11116. (g) Baldauff, E. A.; Buriak, J. M. *Chem. Commun.* **2004**, 2028. (h) Du, H.; Lu, Y.; Yang, W.; Wu, M.; Wang, J.; Zhao, S.; Pan, M.; Cheng, J. *Anal. Chem.* **2004**, *76*, 6166.
(5) Stojanovic, M. N.; Green, E. G.; Semova, S.; Nikic, D. B.; Landry, D. W. *J. Am. Chem. Soc.* **2003**, *125*, 6085.
(6) Kato, T.; Yano, K.; Ikebukuro, K.; Karube, I. *Nucleic Acids Res.* **2000**, *28*, 2000.
(7) Vapnik, V.; Chervonenkis, A. *Theory Probab. Its Appl. (Engl. Transl.)* **1971**, *16*, 264.

10.1021/ja0642663 CCC: $33.50

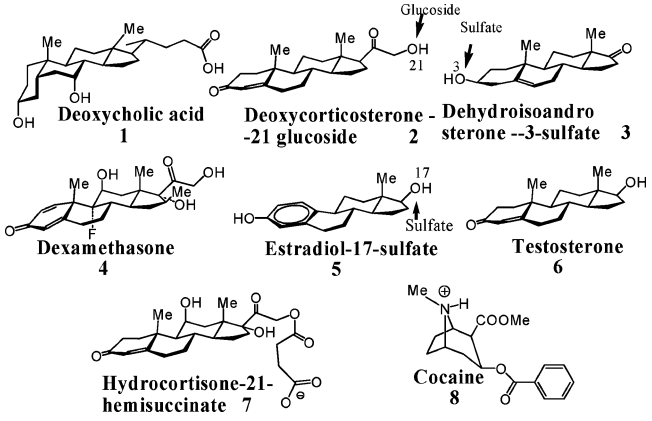

Figure 1. Target steroids and cocaine.

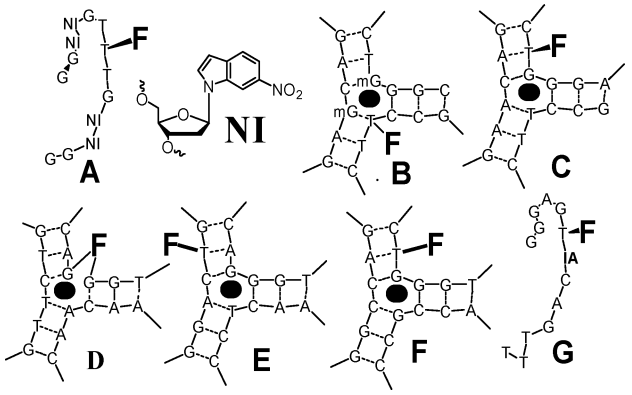

Figure 2. Sensors used in the cross-reactive array construction (full structures given in Materials and Methods). **NI** is nitroindol base, **F** is fluorescein, either attached to T or as a spacer (in **D**), **IA** is spacer with an internal amine, while mG is methyl phosphonate analogue.

genesis, such as errors in bile acid metabolism[8] and congenital adrenal hyperplasia,[9] the concentration of one of the families of steroids increases 1 or 2 orders of magnitude over specimens from healthy subjects. Furthermore, these families of steroids always represent dominant components in mixtures, with high micromolar concentrations. These families are well represented in the screening process by a single steroid with a characteristic set of functionalities (e.g., **1** in Figure 1 for bile acid metabolism errors).

We now report an approach to the selection from a large candidate collection of optimal sensors (Figure 2) to constitute a cross-reactive array with sufficient resolution to differentiate reliably all seven steroids at a range of concentrations. The concentration ranges are limited by the solubility of steroids on the high end, and the sensitivity of the sensors on the low end. We know that some alkaloids, such as cocaine (**8**), cross-react with some three-way junctions,[5] and we included **8** as a control for an "unknown" nonsteroidal interference.

Results and Discussion

Collection of Sensors. Over the past several years, we have assembled a collection of several hundred various fluorescent sensors. They belong to five different types: (1) sensors based on three-way junctions that have to be coupled postsynthesis to fluorophores;[5] (2) sensors based on three-way junctions with fluorophores already incorporated[10] (e.g., **B**−**F** in Figure 2); (3) combinatorial self-assembling sensors;[11] (4) single-chain sensors with four or more hydrophobic base analogues (e.g., nitroindol **NI** in Figure 2, difluorotoluene, nitropyrolle) and their rational modifications (e.g., **A** in Figure 2); and (5) single-chain sensors derived from the GGGAG(fluorescein-dT)TCAGTTT (e.g., **G** in Figure 2) sequence and their rational modifications. Types (4) and (5) were serendipitously discovered during control screening of individual chains of self-assembling sensors.

The detailed study of the mechanism of sensors under (4) and (5) is beyond the scope of the current paper; suffice to say, these sensors are based on dissociation of hydrophobic aggregates. These aggregates are for **A** sufficiently large so that they can be precipitated through centrifugation at high speeds, while hydrophobic molecules inhibit precipitation. In the case of **G**, the aggregates cannot be precipitated, and they are apparently based on stacked G-quartet formation. The incorporation of type (3) sensors did not affect the classification of our steroids; hence they will not be discussed further. In our previous work, modified aptamer precursors for type (1) sensors were transformed into sensors by means of the functionalization of a phoshorothioate, substituting one of the phosphodiester bonds (alternatively amines were used[5]). Because of the presence of diastereomers in most type (1) sensors, and batch-to-batch variations in response to steroids with these sensors, we decided to focus in this work (in contrast to our previous paper) primarily on the three-way junctions with fluorescein already incorporated during oligonucleotide synthesis (type (2)). Our collection of three-way junction sensors contained modifications in sequence, backbone, position, and type of fluorophores, and various spacers (some examples are in Figure 2).

While the steroids used in our studies represented an adequate sample of the entire steroid space, and major variations in steroid structures were covered, there was one notable exception: cholesterol. Despite numerous attempts, we found no sensor responsive to solubilized cholesterol derivatives. We attribute this observation to the self-aggregating properties of cholesterol, even at low concentrations, and the inability of three-way junction sensors to interact with aggregates.

Selection of Sensors for an Array. Our goal is to select a minimal set of sensors, which are able to generate a unique signature for each steroid. We now demonstrate that an empirical procedure for selection is possible, leading to a small set of candidate sensors, which behave as an optimal cross-reactive array. The procedure for selection is as follows: We have a set of compounds that have to be distinguished (Figure 1). To maximize the overall response of sensors, rather than selectivity, let us first select a compound that strongly interacts with most of the sensors (that is, an "anchor" compound). In our case, that compound is deoxycorticosterone (**2**). Next, we choose a potential analyte that is most different from **2**, in our case, deoxycholic acid, **1**. Next, let us define a descriptor that allows an ordering of the steroids with respect to the features distinguishing the two selected compounds (Figure 3A). We name this arbitrary descriptor shape-length, because **1** and **2** are most different in the shape of the hydrophobic core (**1** is bent, **2** is straight) and the length of the side chain (side chain

(8) Yousef, I. M.; Perwaiz, S.; Lamireau, T.; Tuchweber, B. *Med. Sci. Monit.* **2003**, *9*, MT21.
(9) New, M. I. *Mol. Cell. Endocrinol.* **2003**, *211*, 75.
(10) Jhaveri, S. S.; Kirby, R.; Conrad, R.; Maglott, E. J.; Bowser, M.; Kennedy, R. T.; Glick, G.; Ellington, A. D. *J. Am. Chem. Soc.* **2000**, *122*, 2469.
(11) Stojanovic, M. N.; de Prada, P.; Landry, D. W. *J. Am. Chem. Soc.* **2000**, *122*, 11547.

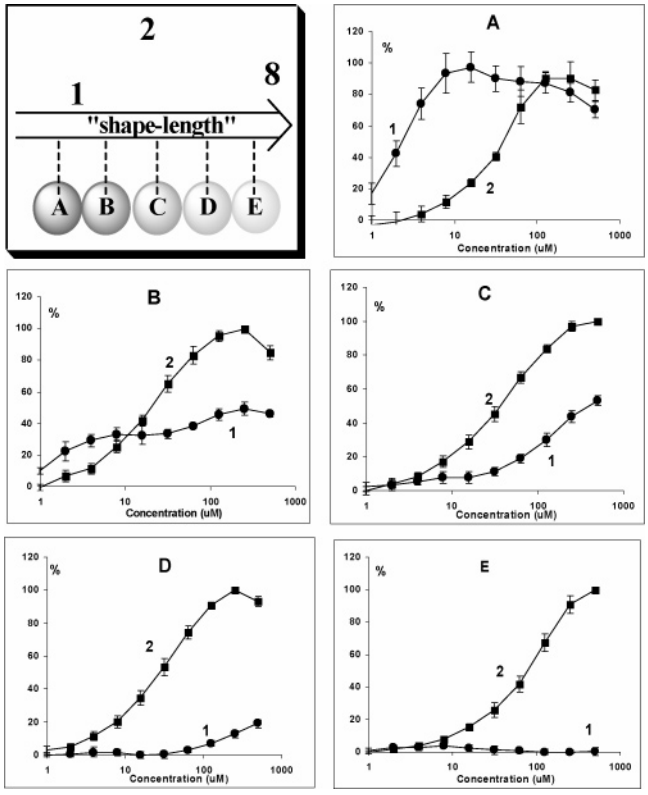

Figure 3. Panel 1 shows the relationship of the shape-length of analytes **1**, **2**, and **8**. By examining the response of each of the sensors **A**−**E** to the steroids **1** and **2**, we then arrive at a ranking for the sensors with respect to shape-length. As shown in the plots, sensor **A** has the greatest relative response to **1** and is thus closest to **1** on the shape-length scale. In contrast, sensor **E** has the least relative response to **1** and is thus furthest from **1** on the scale. Analyte **2** was equidistant from all sensors, while **8**, being reactive mostly with **E**, is on the other extreme of shape length from **1**. The plots show averages of six measurements with standard deviations, and the outputs have been normalized to show percentage increase in fluorescence.

in **1** is much longer than that in **2**). Now, let us find a set of sensors, which are each sensitive to a different region of the shape-length scale. This is done by searching for sensors that show maximal variation in interactions with these two representative compounds.

Such a procedure leads to a selection of cross-reactive sensors **A**−**E** (Figure 2). These sensors give excellent distribution in selectivity along the shape-length parameter as defined by **1** and **2**, indicating the satisfactory richness of motifs in our collection of sensors (at least when it comes to this descriptor). Sensor **A** is selective for **1** over **2** (Figure 3), while, in contrast, sensors **C** and **D** show various degrees of selectivity for **2** over **1**. Sensor **E** was selected because it showed almost no response to **1**, yet it proved cross-reactive with other analytes. Sensor **B** is more responsive to **1** than to **2**, but **1** has lesser efficacy (we define efficacy, as in pharmacology, as the strength of coupling between a recognition process and the downstream signal). Interestingly, sensors such as **B**, with different efficacy toward analytes, can be advantageous in cross-reactive arrays. For example, sensors such as **A** and **B**, with very similar binding selectivity for **1** over **2**, but different efficacy of one of them, would not be considered redundant.

Sensor **B** has a biphasic binding response to deoxycholic acid. This is indicative of two molecules of cholic acid binding. The binding of three-way junctions to more than one steroid has not been reported, and we did not observe it with our other steroid targets, indicating that this behavior is characteristic of a particular sensor−analyte couple (cholic acids vs **B**).

Sensors **G** and **F** have responses similar to those of sensors **A** and **C**, respectively. **G** was different from **A** because it showed slightly improved separation between **2** and **6** (full concentration−response curves for all sensors are available in the Supporting Information), which we considered a critical pair of analytes, due to their very similar structures. **F** was nearly identical to **C**, and we added it to study the influence of redundant sensors on an array. Thus, sensors **A**−**G** formed a somewhat redundant set of initial sensors, from which we wanted to choose the final, optimal set after further measurement and analysis.

All sensors used in this work show excellent reproducibility, indicating that patterns reported in this paper should be easily reproduced in other laboratories. Sensors **A** and **G** show greater sensitivity to experimental conditions (e.g., salt, preheating, incubation, type of test tubes used), which is probably the main reason for **G** not performing optimally in an array (**G** has 3-fold lower efficacy than **A**, which makes the standard error more significant). We note, however, that even an array with only three-way junction sensors has an excellent accuracy in classification (see below).

Determination of Sensitivity. We decided that for an array to be considered responsive to a certain steroid concentration, at least one of the sensors must show a change in fluorescence of 10% or more. On the basis of this criterion, we selected the following ranges of concentrations for further study: **1**, 2−500 μM; **2**, 8−500 μM; **3**, 32−500 μM; **4**, 50 and 100 μM; **5**, 32−125 μM; **6**, 5−150 μM; **7**, 125−500 μM; **8**, 32−500 μM. In the case of **4**, **5**, and **6**, the maximum concentration is less than 500 μM because the solubility of the steroid is an issue at higher concentrations. Tests used for presentation here were run in two sets of triplicates, on two different days within 1 month.

Classification and Visualization Methods. We wanted to show that the array is able to uniquely identify any given steroid, at any concentration, within the bounds of sensitivity. Once an unknown is classified, regression can be used reliably to estimate concentration. Thus, for each of the eight analytes, we formed a class consisting of all measurements for all concentrations of the analyte, for which the array showed sensitivity. Each class represents the range of possible signatures that an analyte is expected to give. As long as these classes do not overlap much, the array will have the potential to uniquely identify the analytes.

Because each of our classes consists of points from several different concentration values, the classes do not exhibit any well-known statistical distribution; thus, we chose to implement a *K*-nearest-neighbors (KNN) classifier.[12,13] This is a nonparametric technique, which does not attempt to model the data as a distribution. Instead, it chooses a classification for an unknown value by taking a vote of the identities of the *K*-nearest neighbors in the *N*-dimensional space (where *N* is the number of sensors in an array). The class chosen is the one with the most representatives in the set of nearest neighbors. We empirically settled on $K = 3$ and broke ties according to the closest point.

To understand if the array is useful for identifying steroids, we must obtain an estimate of its classification accuracy. We

(12) Scholkopf, B.; Smola, A. J. *Learning with Kernels*; 2002.
(13) Duda, R. O.; Hart, P. E. *Pattern Classification*; 2001.

Table 1. Top Ten Sensor Combinations, and 95% Confidence Intervals for Their True Error, As Estimated by the LOO Method

A	B	C	D	E	F	G	accuracy
X	X	X	X	X	X		96.1% ± 2.3%
X	X	X		X	X		95.3% ± 2.5%
X	X	X	X	X	X	X	94.5% ± 2.7%
X		X	X	X	X		94.5% ± 2.7%
X	X	X		X	X		94.1% ± 2.8%
X	X	X		X			93.7% ± 2.9%
X	X	X	X	X		X	93.7% ± 2.9%
	X	X	X	X	X		93.3% ± 3.0%
X	X			X	X		92.8% ± 3.1%
X	X	X	X	X			92.8% ± 3.1%

applied the leave-one-out (LOO) or jackknife estimator to our classifier, because LOO is an unbiased method and does not require any additional data for an estimation of an error rate. All error rates reported are estimated with the 95% confidence level.

Using all seven sensors, we obtained an accuracy rate of 94.5% ± 2.7%. We are interested in minimizing the number of sensors needed in the array, so we examined the LOO accuracy rate of the 127 arrays consisting of every possible subset of the original seven sensors. Of the top 10 sensor combinations listed in Table 1, the combination that omits sensor **G** has the best overall performance. While the difference from the array consisting of all seven sensors is not statistically significant (and, as a matter of fact, the difference between the first 10 arrays is not statistically significant at 95% confidence level), it is still preferable to reduce the number of sensors. Therefore, the recommendation is that the final array should consist of six sensors, excluding **G**. Because of the unique ability of **E** to effectively eliminate the whole class of 1, **E** was the only sensors selected in all top 10 subarrays.

While studying some of the other candidates for the optimal sensor set, we made some interesting observations. For example, the LOO accuracy rate for an array consisting of only five sensors **A**−**E** is about 93%. To determine which classes are creating a problem for these sensors, we generated a confusion matrix (Supporting Information). This protocol records the classification choice the classifier has made for each of the training points when it is treated as an unknown. With the initial five sensors, **A**−**E**, 16 out of a total of 240 samples were misclassified, with 6 of the errors involving confusion between **5**, at low concentrations, and **8**. Our array with the best result (96% accuracy) adds one redundant sensor **F**, reducing the number of total errors to eight, and the number of errors between **5** and **8** to only one. This observation could be rationalized by studying individual dose−response curves for sensors **C** and **F** (Supporting Information); estradiol and cocaine are very close in these two sensors, and they are only somewhat better separated by **F**. Thus, introduction of redundancy had two effects: (1) it effectively increased the number of measurements in this area; and (2) it gave a slight improvement of resolution at the lowest concentrations of estradiol; cumulatively, these two changes seemed to have eliminated some confusion in our array.

If we focused on only the three-way junction derived sensors, we would also obtain correct classification in approximately 93% of cases, but with different sources of most common classification errors: 6 out of 15 errors are coming from misclassifying **6** as **2** and vice versa (with further errors misclassifying **2** and **3**), and only one misclassification of **5** and **8** (the major source of misclassification with the initial five sensors). This kind of misclassification, when compared to the six-sensor array (or any subarray containing **A**-type sensors), indicates that three-way junctions excel in the recognition of hydrophobic core, while side chains are better distinguished by sensors such as **A**. This is in agreement with the proposed mechanism of action of both classes of sensors, in which displacement of the fluorophore embedded in the three-way junction leads to sensing, while even a straight chain can help in inhibition of the assembly of nitroindole rings.

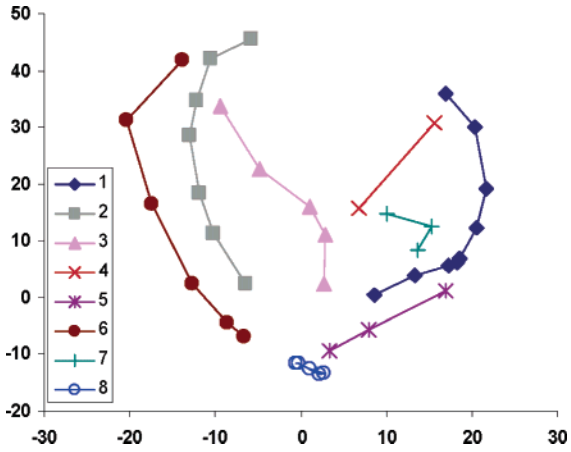

Figure 4. Projections from the seven-dimensional space onto a plane for selected concentrations of each of the eight (including cocaine) compounds. The values assigned to points are given as: $x = 0.190(\mathbf{E}) + 0.312(\mathbf{B}) + -0.0504(\mathbf{C}) + 0.128(\mathbf{A}) + -0.0940(\mathbf{G}) + -0.861(\mathbf{D}) + 0.313(\mathbf{F})$, and $y = -0.122(\mathbf{E}) + -0.263(\mathbf{B}) + 0.919(\mathbf{C}) + -0.00160(\mathbf{A}) + 0.192(\mathbf{G}) + -0.185(\mathbf{D}) + 0.00192(\mathbf{F})$.

To provide an overall impression of the relation between the classes (analytes) in *N*-dimensional space, some type of a 2D representation is preferred, because it takes advantage of the natural human ability for visual pattern recognition. Unfortunately, we found that for our data the restriction to two dimensions imposed too severe a constraint on standard methods, such as multidimensional scaling,[14] while shape and distribution of the classes precluded using principal components analysis[15] (PCA) and multiple discriminate analysis[13] (MDA). However, it is possible to derive a projection, which separates the classes well by directly optimizing for this quality. Given that each class can be approximated by a curve connecting the average responses for each concentration, we developed a new projection pursuit[16] algorithm, based on finding the projection maximizing the separation in the plane of all seven curves. The optimal projection, shown in Figure 4, gives a visual impression of the relationships among the classes. This representation is a 2D projection from the 7D space consisting of all of the sensors in the test set. This projection, while it gives us an impression of separation in seven-dimensional space, is a separate and unrelated analysis of the data from the classification accuracy. Thus, this method does not provide an insight into the exact areas of misclassification, for which one needs to look at the confusion matrixes to understand (Supporting Information).

An interesting insight into the seven-dimensional space obtained with these sensors can be gained by analyzing our

(14) Cox, T. F.; Cox, M. A. A. *Multidimensional Scaling*; 2001.
(15) Jolliffe, I. T. *Principal Component Analysis*; 1986.
(16) Huber, P. J. *Ann. Stat.* **1985**, *13*, 435−475.

optimized projection in Figure 4. The algorithm used maximizes the distance between individual curves (classes), and it clustered together compounds with aromatic groups (**5** and **8**, bottom right of Figure 4), compounds with no, or one, hydroxyl group within the steroid core (**2**, **3**, and **6**, upper left of Figure 4), and compounds with two hydroxyl groups (**1**, **4**, and **7**, upper right of Figure 4). This kind of clustering can actually be intuitively expected in the optimized projections, as related compounds with similar overall sensor responses should be located closer in the plot. For example, as mentioned above, **5** and **8** form a group and they both have particularly strong interactions with **E**, which would separate them from the rest of the analytes. Also, the group consisting of **2**, **3**, and **6** has stronger interactions with all negatively charged three-way junctions than do other steroids.

Conclusions

Through screening pure steroids of various structures with our collection of sensors, we can increase our understanding of the "synthetic" or "bottom-up" approach to the construction of cross-reactive arrays. We expect that the methodology used to select sensors in this paper is useful as a general selection procedure for small chemical sensor arrays. The method used here applies screening to select a somewhat redundant set of sensors, which is then expected to cover the entire space of possible analytes. This initial set of sensors should be kept small to limit the number of measurements required, but must contain enough sensors to provide adequate discriminatory power. Once this initial set is formulated, and its response to each of the analytes is measured, an optimal subset can be found by deciding on an appropriate classifier, and analyzing the error rates of classifiers formed from every possible subset of sensors. Such arrays would be particularly useful in differential diagnosis of errors in steroidal metabolism that lead to gross increases in concentration of a family of metabolites. The arrays can operate over a range of analyte concentrations, which makes them very useful in urinalysis, where differences in kidney filtrations can complicate analysis, leading to variability in concentrations, and necessitating standardization against creatinine.

The success of our cross-reactive array in classifying the more general set of steroids demonstrates the usefulness of DNA three-way junction-based receptors as scaffolds for the construction of sensors. This particular scaffold has been identified from the corresponding sets of aptamers,[6] and it is intriguing to suggest that a similar approach, combining the power of SELEX[18] for the initial scaffold identification with a rational/organic-synthetic combinatorial approach for diversification, could be used to identify further scaffolds suitable for other classes of compounds of medical interest, in particular, oligosaccharides and lipids.

Materials and Methods

Materials. Sensors **A**, **C**, **E**, **F**, and **G** were custom-made and HPLC purified by Integrated DNA Technologies, Inc. (Coralville, IA) and were used as received. Sensor **B** and the three-way junction with two positive charges in the Supporting Information were custom-made and PAGE purified by Trilink Biotechnologies (Carlsbad, CA). Sensor **D** was made in house, using phosphoramidites and other chemicals from Glen Research (Sterling, Virginia) according to manufacturer procedures on 8909 DNA RNA synthesizer (Applied Biosystems, Foster City, CA), followed by PAGE purification. DNAse/RNAse free water was purchased from ICN (Costa Mesa, CA) and used for all buffers, and for stock solutions of oligonucleotides. **A**: 5′ GG(NI)(NI)GT(fT)TG-(NI)(NI)GG; NI-nitroindole base. **B**: ATA TGA CmGA GGA TAA ATC CT(fT) CCG CGA AGC GGmG TCA TAT; m, methyl phosphonate. **C**: GGGAGACAAGGATAAATCCTTCCGCGAAGAGGG(fT) CGACA; fT, fluoresceinated dU base. **D**: GGGAGTCAGGATAAATC-CTCAACGAAGTGG(spF) GACGACA; spF, spacer fluorescein. **E**: GGGAG(fT)CAGGATAAATCCTCAACGAAGTGGGACGA CA. **F**: ATATGACCGGATAAATCCGCCACGAAGTGGG(fT) CATAT. **G**: GGGAG(fT)(intNH_2)CAGTTTTTTTT; intNH_2, spacer with amino group.

Instruments. Fluorescent spectra were taken on a Perkin-Elmer (San Jose, CA) LS-55 luminometer with a Hamamatsu Xenon lamp. Experiments were performed at the excitation wavelength of 480 nm and emission scan of 500−700 nm. Array work was done on a Perkin-Elmer Victor II microplate reader, using filters of appropriate wavelengths. The gels were photographed using AlphaImager 3400 (Alpha Innotech).

Measurements. In the initial screening of all sensors, 1 and 2 were used (with various other steroids) at concentrations from 1 to 2000 μM and 1 to 1000 μM, respectively, in a "selection buffer", as reported previously.[6] For the purpose of final array characterization, samples of each analyte were run in triplicates at each concentration and for each sensor, on 2 days, for a total of six measurements for each concentration and each sensor. Sensors **B**−**F** were run at 10 nM concentrations, while **A** and **G** were run at 40 nM concentrations, all in "selection buffer", after preheating and cooling. In contrast to **B**−**F**, readings from **A** and **G** were time-sensitive. Thus, all readings were performed after the signals for **A** and **G** stabilized, that is, after 3 h. Readings were standardized, with the highest reading for any given sensor set on 100%, and buffer average at 0%. In absolute values, the maximum increase in fluorescence was for each sensor as follows: **A** 200%, **B** 80%, **C** 120%, **D** 120%, **E** 120%, **F** 100%, **G** 95%.

Acknowledgment. This paper is dedicated to Prof. Yoshito Kishi on the occasion of his 70th birthday. We are grateful to Kay Severin and Sandra Bencic (from David R. Walt's group at Tufts University) for advice on using various existing methods for characterization of cross-reactive arrays. The funding by the NSF (Biophotonics, BES-032197) is gratefully acknowledged. Initial collection of sensors was constructed with help of funding by the NIH (NIBIB) and NASA. M.N.S. is a Searle Scholar.

Supporting Information Available: Confusion matrixes, a comment on 2D projections, and dose−response curves used to analyze our collections of sensors in arrays for all sensors and all steroids. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0642663

(17) Knemeyer, J.-P.; Marne, N.; Sauer, M. *Anal. Chem.* **2000**, *72*, 3717−3724.
(18) Wilson, D.; Szostack, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611−647.