J|A|C|S
ARTICLES

Published on Web 08/19/2009

Pattern-Based Recognition for the Rapid Determination of Identity, Concentration, and Enantiomeric Excess of Subtly Different Threo Diols

Shagufta H. Shabbir, Leo A. Joyce, Gabriella M. da Cruz, Vincent M. Lynch, Steven Sorey, and Eric V. Anslyn*

Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78712

Received June 4, 2009; E-mail: anslyn@austin.utexas.edu

Abstract: A pattern-based recognition approach for the rapid determination of the identity, concentration, and enantiomeric excess of chiral vicinal diols, specifically threo diols, has been developed. A diverse enantioselective sensor array was generated using three chiral boronic acid receptors and three pH indicators. The optical response produced by the sensor array was analyzed by two pattern-recognition algorithms: principal component analysis and artificial neural networks. Principal component analysis demonstrated good chemoselective and enantioselective separation of the analytes, and an artificial neural network was used to accurately determine the concentrations and enantiomeric excesses of five unknown samples with an average absolute error of ±0.08 mM in concentration and 3.6% in enantiomeric excess. The speed of the analysis was enhanced by using a 96-well plate format, portending applications in high-throughput screening for asymmetric-catalyst discovery. X-ray crystallography and ^{11}B NMR spectroscopy was utilized to study the enantioselective nature of the boronic acid host **2**.

Introduction

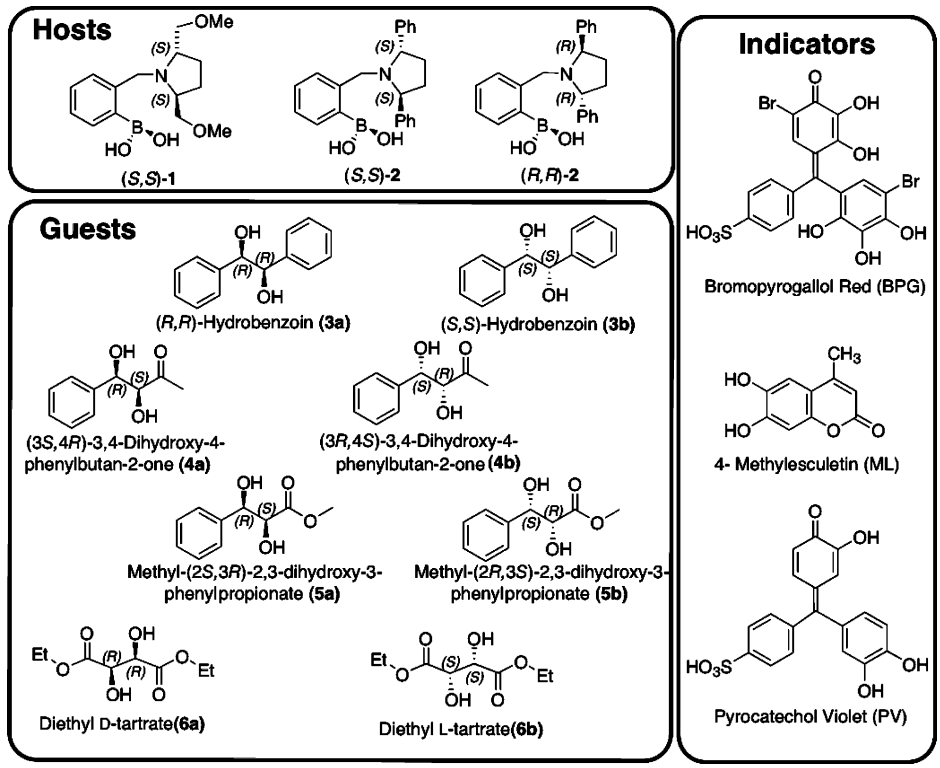
Animals are constantly surveying their external environment via their olfactory system for chemicals that indicate food sources and habitats as well as chemical signals controlling social interactions and reproductive behavior.[1] The olfactory system involves a recognition protocol where the detectors/sensors generate a pattern that is recognized and classified by the brain.[2] Inspired by the olfactory system, the use of multivariate data analysis combined with sensors of partially overlapping selectivity has become a powerful tool in the field of molecular recognition.[3−21]

In an analogous fashion, this report describes a novel optical recognition scheme for the rapid determination of the identity, enantiomeric excess (*ee*), and concentration ($[G]_t$) of chiral vicinal diols, particularly threo diols having close structural similarity. The differential receptors are chiral boronic acids, which are responsible for the sensors' chemoselective and enantioselective discriminatory properties. The signaling relies upon indicator displacement assays (IDAs) to provide a colorimetric response. The assay was developed by using combinations of three differential receptors and three different indicators within a single array, which resulted in fingerprinting of each diol analyzed.

Indicator displacement assays rely on colorimetric or fluorescence indicators that reversibly interact with a receptor and, upon binding to the receptor, undergo changes in their optical properties. A competition between the analyte and the indicator for the binding site of the receptor generates an equilibrium in solution that can be monitored optically and related to $[G]_t$.[22]

(1) Breer, H.; Fleischer, J.; Strotmann, J. *Cell. Mol. Life Sci.* **2006**, *63*, 1465–1475.
(2) Doty, R. L. *Handbook of Olfaction and Gustation*, 2nd ed.; Marcel Dekker: New York, 2003.
(3) Janes, L. E.; Lowendahl, A. C.; Kazlauskas, R. J. *Chem.—Eur. J.* **1998**, *4*, 2324–2331.
(4) Badalassi, F.; Klein, G.; Crotti, P.; Reymond, J.-L. *Eur. J. Org. Chem.* **2004**, 2557–2566.
(5) Ruta, J.; Ravelet, C.; Baussanne, I.; Decout, J.-L.; Peyrin, E. *Anal. Chem.* **2007**, *79*, 4716–4719.
(6) Simeonov, A.; Matsushita, M.; Juban, E. A.; Thompson, E. H. Z.; Hoffman, T. Z.; Beuscher, A. E., IV; Taylor, M. J.; Wirsching, P.; Rettig, W.; McCusker, J. K.; Stevens, R. C.; Millar, D. P.; Schultz, P. G.; Lerner, R. A.; Janda, K. D. *Science* **2000**, *290*, 307.
(7) Johnson, S. R.; Sutter, J. M.; Engelhardt, H. L.; Jurs, P. C.; White, J.; Kauer, J. S.; Dickinson, T. A.; Walt, D. R. *Anal. Chem.* **1997**, *69*, 4641–4648.
(8) Rakow, N. A.; Suslick, K. S. *Nature* **2000**, *406*, 710.
(9) Lavigne, J. J.; Anslyn, E. V. *Angew. Chem., Int. Ed.* **2001**, *40*, 3118–3130.
(10) Wiskur, S. L.; Floriano, P. N.; Anslyn, E. V.; McDevitt, J. T. *Angew. Chem., Int. Ed.* **2003**, *42*, 2070–2072.
(11) Zhou, H.; Baldini, L.; Hong, J.; Wilson, A. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2006**, *128*, 2421–2425.
(12) Palacios, M. A.; Nishiyabu, R.; Marquez, M.; Anzenbacher, P., Jr. *J. Am. Chem. Soc.* **2007**, *129*, 7538–7544.
(13) Nieto, S.; Lynch, V. M.; Anslyn, E. V.; Kim, H.; Chin, J. *J. Am. Chem. Soc.* **2008**, *130*, 9232–9233.
(14) Ponnu, A.; Edwards, N. Y.; Anslyn, E. V. *New J. Chem.* **2008**, *32*, 848–855.
(15) Reetz, M. T.; Zonta, A.; Schimossek, K.; Leibeton, K.; Jaeger, K.-E. *Angew. Chem., Int. Ed.* **1997**, *36*, 2830–2832.
(16) Matsushita, M.; Yoshida, K.; Yamamoto, N.; Wirsching, P.; Lerner, R. A.; Janda, K. D. *Angew. Chem., Int. Ed.* **2003**, *42*, 5984–5987.
(17) Dey, S.; Karukurichi, K. R.; Shen, W.; Berkowitz, D. B. *J. Am. Chem. Soc.* **2005**, *127*, 8610–8611.
(18) Abato, P.; Seto, C. T. *J. Am. Chem. Soc.* **2001**, *123*, 9206–9207.
(19) Li, Z.-B.; Lin, J.; Qin, Y.-C.; Pu, L. *Org. Lett.* **2005**, *7*, 3441–3444.
(20) Eelkema, R.; van Delden, R. A.; Feringa, B. L. *Angew. Chem., Int. Ed.* **2004**, *43*, 5013–5016.
(21) Taran, F.; Gauchet, C.; Mohar, B.; Meunier, S.; Valleix, A.; Renard, P. Y.; Creminon, C.; Grassi, J.; Wagner, A.; Mioskowski, C. *Angew. Chem., Int. Ed.* **2002**, *41*, 124–127.

Scheme 1. Structures of Hosts, Guests/Analytes, and Indicators Used in This Study



Our group has pioneered the development of enantioselective IDAs that utilize a chiral receptor and in which the color change is related to *ee* and $[G]_t$.[23,24]

The data generated by the sensor arrays is commonly processed by unsupervised and supervised pattern-recognition algorithms. A major unsupervised technique is principal component analysis (PCA), while an artificial neural network (ANN) is a supervised technique. PCA reduces multidimensional and partly correlated data to two, three, or more dimensions. This is achieved by projecting the data onto fewer dimensions that represent maximum variance relationships between variables.[25] ANN programs are based on a simplified model of the brain. ANN analysis requires a training set that consists of a collection of known parameters. The training set creates a neural network suitable for the analysis at hand. This network is then used to analyze unknown samples.[26,27]

The goal of the present study was to demonstrate the ability to distinguish the chiralities as well as the chemical identities of subtly different analytes; therefore, diols **3**−**6** were selected as targets (Scheme 1). Analyte **3** has two phenyl groups, in contrast to **4** and **5**, which contain only one. Analytes **4** and **5** differ through the inclusion of a ketone as opposed to an ester, respectively. Lastly, analyte **6** contains two esters and no phenyl groups. Hence, we challenged our array approach by using four analytes and their enantiomers possessing subtly different electron-withdrawing groups.

Our previous studies have shown that the use of cross-reactive sensors in a single array enhances the fingerprinting of the analytes, allowing one to determine structural similarities and chirality simultaneously.[28] Furthermore, increasing the enantioselectivity of the receptors improves the chiral discrimination. On the basis of previous results showing that C_2-symmetric chiral secondary amines on the boronic acid give excellent enantioselectivity, host **1** was chosen.[29] With this guiding principle, we also introduced host **2**, which is also C_2-symmetric about the pyrrolidine ring (Scheme 1). The aim was to develop a second highly enantioselective host that can be used in conjunction with host **1** in our pattern-recognition protocol.

The use of an additional chiral receptor could allow for concentration as well as *ee* determination. In our previous studies, an achiral receptor was consistently used to determine concentration.[30] In the senses of taste and smell, however, the suites of receptors used are all chiral by virtue of the structure of their proteins. Nevertheless, both handedness and relative concentrations of tastants and odorants can be determined. Hence, another goal of the work described herein was to show that a suite of chiral synthetic receptors could do the same. Furthermore, in this study we explored the reasons for the enantioselectivities of the host **2** via X-ray crystallography and ^{11}B NMR spectroscopy.

Results and Discussion

Synthesis. Hosts **1** and **2** were synthesized by reductive amination according to a published literature procedure.[31] The

(22) Nguyen, B. T.; Anslyn, E. V. *Coord. Chem. Rev.* **2006**, *250*, 3118–3127.
(23) Zhu, L.; Shabbir, S. H.; Anslyn, E. V. *Chem.—Eur. J.* **2007**, *13*, 99–104.
(24) Leung, D.; Anslyn, E. V. *J. Am. Chem. Soc.* **2008**, *130*, 12328–12333.
(25) Jurs, P. C.; Bakken, G. A.; McClelland, H. E. *Chem. Rev.* **2000**, *100*, 2649–2678.
(26) Jansson, P. A. *Anal. Chem.* **1991**, *63*, 357A–362A.
(27) Burns, J. A.; Whitesides, G. M. *Chem. Rev.* **1993**, *93*, 2583–2601.
(28) Folmer-Andersen, J. F.; Kitamura, M.; Anslyn, E. V. *J. Am. Chem. Soc.* **2006**, *128*, 5652–5653.
(29) Zhu, L.; Zhong, Z.; Anslyn, E. V. *J. Am. Chem. Soc.* **2005**, *127*, 4260–4269.
(30) Shabbir, S. H.; Regan, C. J.; Anslyn, E. V. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10487–10492.

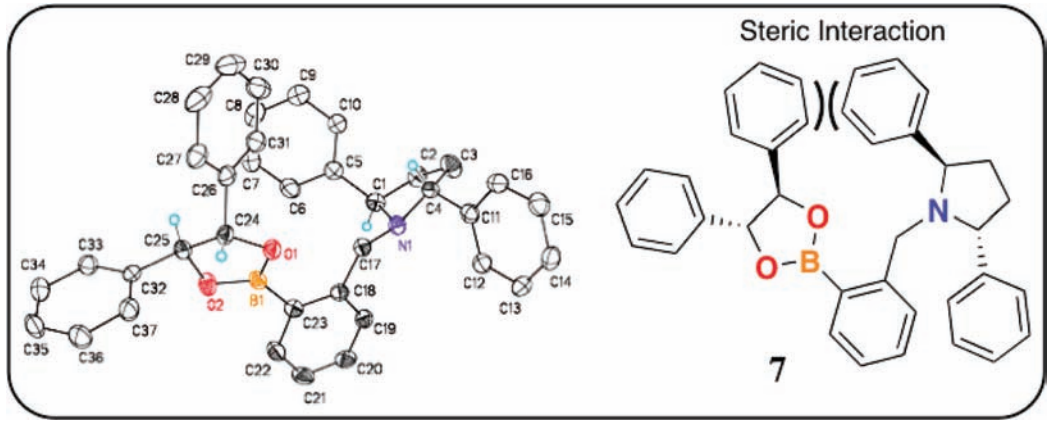


Figure 1. (left) Crystal structure of **7**, the complex between host (*R*,*R*)-**2** and (*S*,*S*)-hydrobenzoin (**3b**). (right) ChemDraw representation of the crystal structure of **7**.

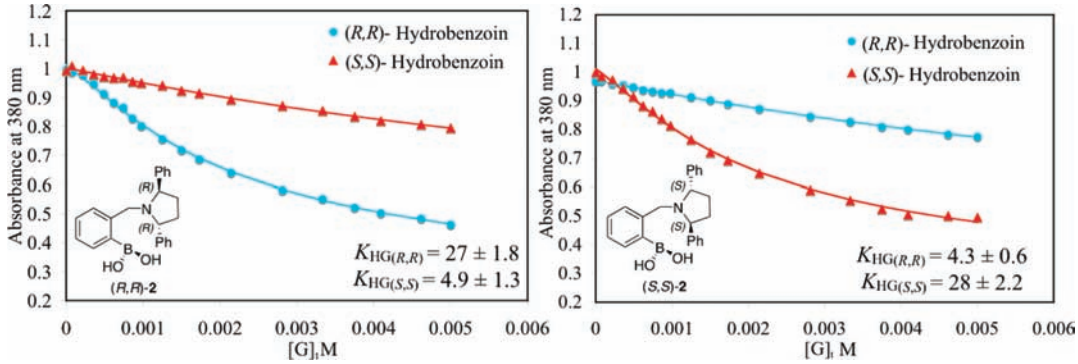


Figure 2. Enantioselective indicator displacement assays of hosts (left) (*R*,*R*)-**2** and (right) (*S*,*S*)-**2** with guests (*R*,*R*)-hydrobenzoin (**3a**) and (*S*,*S*)-hydrobenzoin (**3b**), using ML as the indicator. All of the titrations were carried out with 440 μM **2** and 75 μM ML in a 10 mM *p*-toluenesulfonic acid/Hunig's base buffer (pH 7.4) in 100% MeOH. Values of the association constants K_{HG} are given in units of 10^2 M^{-1}. All of the measurements were taken at 25 °C. The solid lines are calculated curves resulting from iterative data fitting for a displacement assay.[38]

diols used in this analysis are commercially available except for **4**, which was synthesized by a previously established Sharpless asymmetric dihydroxylation reaction.[32] Diol (3*S*,4*R*)-3,4-dihydroxy-4-phenylbutan-2-one (**4a**) was isolated in 70% yield and 93% *ee*, while (3*R*,4*S*)-3,4-dihydroxy-4-phenylbutan-2-one (**4b**) was obtained in 76% yield and 86% *ee*. The *ee*'s of **4a** and **4b** were determined using the NMR analysis method of James and co-workers[33] (Figure S7 in the Supporting Information).

X-ray Crystallographic Analysis. To study the enantioselective behavior of our boronic acid hosts, an X-ray-quality crystal of **7**, the complex between host (*R*,*R*)-**2** and (*S*,*S*)-hydrobenzoin (**3b**), was isolated by diffusing methanol into a concentrated solution of the complex in dichloromethane (Figure 1). Previously, our group analyzed the free boronic acid host *o*-(pyrrolidin-1-ylmethyl)phenylboronic acid and its complexes with catechol by X-ray crystallography.[34] It was shown that in these complexes, the presence of the N−B bond is solvent-dependent. An aprotic solvent ($CHCl_3$) facilitates the formation of the N−B bond, while in a protic solvent (CH_3OH) a solvolysis pathway is operative. When complex **7** was crystallized from methanol, instead of the expected solvent-inserted structure, the boron atom was found to be sp^2 hybridized. The average B−O bond length is 1.36 Å, and the O−B−O angle of the cyclic (*S*,*S*)-hydrobenzoin boronate is 113.23°. As expected in such a case, there is no intramolecular B−N bond; instead, an N−B distance of 4.43 Å was observed. The optimal distance for an N−B bond is 1.5−1.8 Å, as reported for other boronic acid molecules with similar configurations.[35,36] The absence of the solvent-inserted structure could be due to the more sterically inaccessible nitrogen lone pair in the tertiary amine of host (*R*,*R*)-**2**, as we recently found in another sterically congested system.[37] Interestingly, the ^{11}B NMR profile of the complex in CD_3OD indicated that the boron atom is solvated in solution (see the next section). Therefore, there is a discrepancy between the structure of the complex in the solid state and in solution.

Nevertheless, the crystal structure does lend insight into the enantioselectivity of host **2**. The indicator displacement isotherms for hosts (*R*,*R*)-**2** and (*S*,*S*)-**2** with guests **3a** and **3b** using 4-methylesculetin (ML) (Scheme 1) as the indicator are shown in Figure 2. As required by the principal of chirality, the enantiomeric hosts show opposite enantioselectivity for enantiomeric guests. As shown, host (*R*,*R*)-**2** has a lower affinity for **3b** than for **3a**. This can be explained in terms of the crystal

(31) Zhu, L.; Anslyn, E. V. *J. Am. Chem. Soc.* **2004**, *126*, 3676–3677.
(32) Walsh, P. J.; Sharpless, K. B. *Synlett* **1993**, 605–606.
(33) Kelly, A. M.; Perez-Fuertes, Y.; Arimori, S.; Bull, S. D.; James, T. D. *Org. Lett.* **2006**, *8*, 1971–1974.
(34) Zhu, L.; Shabbir, S. H.; Gray, M.; Lynch, V. M.; Sorey, S.; Anslyn, E. V. *J. Am. Chem. Soc.* **2006**, *128*, 1222–1232.
(35) Wiskur, S. L.; Lavigne, J. J.; Ait-Haddou, H.; Lynch, V. M.; Chiu, Y. H.; Canary, J. W.; Anslyn, E. V. *Org. Lett.* **2001**, *3*, 1311–1314.
(36) Wiskur, S. L.; Lavigne, J. J.; Matzger, A.; Tobey, S. L.; Lynch, V. M.; Anslyn, E. V. *Chem.—Eur. J.* **2004**, *10*, 3792–3804.
(37) Collins, B. E.; Sorey, S.; Hargrove, A. E.; Shabbir, S. H.; Lynch, V. M.; Anslyn, E. V. *J. Org. Chem.* **2009**, *74*, 4055–4060.

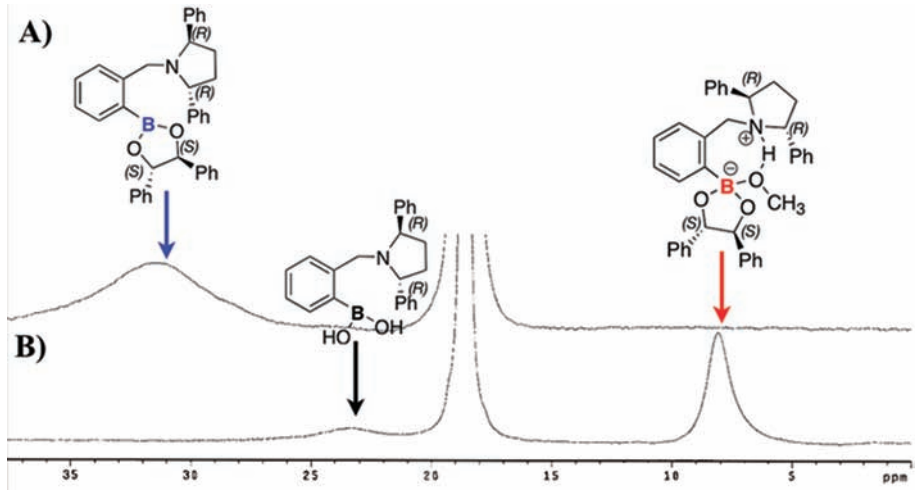


Figure 3. ^{11}B NMR spectra of complex **7** in (A) $CDCl_3$ and (B) CD_3OD. The signal at 18.6 ppm is from the trimethyl borate external standard.

structure, in which the phenyl group on the pyrrolidine ring of (*R*,*R*)-**2** has an unfavorable steric interaction with the phenyl group of **3b**. This unfavorable interaction would not be present in the complex between (*R*,*R*)-**2** and **3a**.

Structural Characterization in Solution Using ^{11}B NMR Spectroscopy. Structural studies using ^{11}B NMR spectroscopy were carried out to investigate the effect of the solvent on the nature of the N−B interaction in complex **7** in solution. The ^{11}B NMR chemical shift is dependent on hybridization, charge, coordination, and the substitutents present on the boron. Therefore, useful information about the structure of the molecule can be obtained by ^{11}B NMR analysis. The ^{11}B NMR signals of the complex and the receptor were assigned on the basis of a comparison with our previous analysis of the *o*-(*N*,*N*-dialkylaminomethyl)arylboronate system.[34]

The sample of crystals used to obtain the X-ray structure of complex **7** displayed one ^{11}B signal at 31.3 ppm in $CDCl_3$ (Figure 3A). The signal at 31.3 ppm is due to the trigonal-planar boron species with no N−B bond. However, in CD_3OD (Figure 3B), the same set of crystals of complex **7** showed two peaks. The major peak (∼90%) at 8.1 ppm is due to the fully solvated sp^3-hybridized boron species of complex **7** in solution, while the minor signal (∼20%) at 23.3 ppm corresponds to the ^{11}B NMR signal of the free host **2** in methanol (see the Supporting Information).

Host **2** by itself has a chemical shift of 28.8 ppm in $CDCl_3$ and 23.3 ppm in CD_3OD. The signal at 28.8 ppm was assigned to the trigonal planar boron in $CDCl_3$. The signal at 23.3 ppm in CD_3OD could not be assigned to a single boron species in solution because, as we have previously shown for the *o*-(*N*,*N*-dialkylaminomethyl)arylboronate system, the N−B-coordinated species has a chemical shift of ∼14 ppm. We postulate that there is a fast exchange between the solvent-inserted sp^3-hybridized boron and the sp^2-hybridized boron of host **2** in methanol. Therefore, the signal at 23.3 ppm could not be assigned to a single structure of host **2** in methanol.

As expected, complex **7** exists as the B-sp^3-hybridized solvated species in CD_3OD with a chemical shift of 8.1 ppm. However, the crystal structure shows the presence of the B-sp^2-hybridized form of complex **7**. In order to explain this discrepancy, we hypothesize that there is an equilibrium between the solvent-inserted sp^3-hybridized boron and the sp^2-hybridized boron of complex **7** in solution but that the sp^2-hybridized species of complex **7** is what crystallizes out of solution.

Table 1. Binding Constants K_{HI} of Host (*S*,*S*)-**2** with the Indicators

indicator	K_{HI} (10^4 M^{-1})
A	8.42 ± 1.76
AC	6.41 ± 2.21
BPG	191 ± 131
ML	3.61 ± 0.13
PG	1.02 ± 0.07
PV	1.36 ± 0.02

In summary, we have studied for the first time the origin of enantioselectivity in our boronic acid-based receptors, and a clear steric clash is seen in the unfavorable complex (Figure 1). Insight into the structure of complex **7** in different solvent systems was also achieved. Complex **7** is B-sp^2-hybridized in $CDCl_3$ while there is an equilibrium between the sp^3-hybridized solvent-inserted boron and the sp^2-hybridized boron in CD_3OD.

Generation of a Chemo- and Enantioselective Array. Before creating our pattern-based recognition procedures, we needed to establish the proper pairings of hosts and indicators to best enantiodifferentiate the chiral diol targets, as described below. However, we first determined binding constants of the hosts with the indicators, because we have found that the optical response of an enantioselective IDA is best when the host is ∼90% saturated with the indicator.[29] Six catechol-based indicators were used in the study described herein: alizarin (A), alizarin complexone dihydrate (AC), bromopyrogallol red (BPG), 4-methylesculetin (ML), pyrogallol red (PG), and pyrocatechol violet (PV). Host **1** was studied previously, and thus, the ratio for achieving 90% saturation of host **1** with each of the six indicators was calculated from the literature binding constants.[30] The binding constants for the indicators with host **2** are given in Table 1 (see Figures S1−S6 in the Supporting Information).

To establish the optimal host−indicator pairings for enantiodiscrimination, we screened host−indicator duos for colorimetric differences between the diol analytes. The concentration of each host that yields 90% saturation with each indicator was used in a 96-well plate (Table S1 in the Supporting Information). The receptor−indicator pairs were treated with the diols (Figure 4) using a diol concentration of 5 mM. The results from these plates identified the three host−indicator combinations **1**−PV, **2**−PV, and **2**−ML as good pairings for enantiomeric discrimi-

(38) Connors, K. A. *Binding Constants: The Measurement of Molecular Complex Stability*; John Wiley and Sons: New York, 1987.

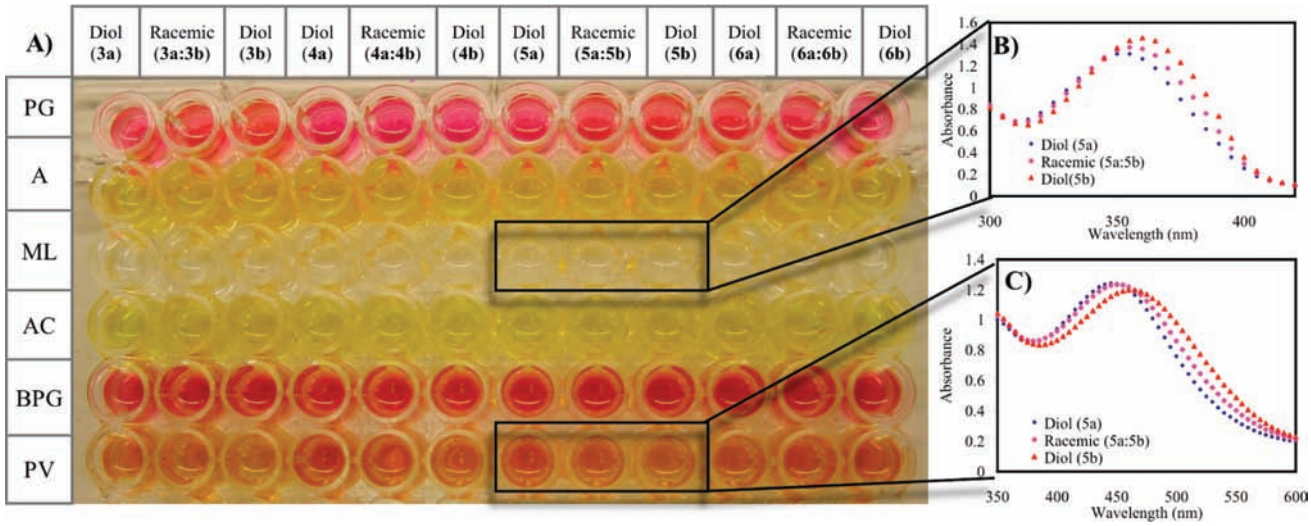


Figure 4. (A) Method used to screen for host (*R*,*R*)-**2** with the indicators pyrogallol red (PG), alizarin (A), 4-methylesculetin (ML), alizarin complexone dihydrate (AC), bromopyrogallol red (BPG), and pyrocatechol violet (PV) against the diols (*R*,*R*)-hydrobenzoin (**3a**), (*S*,*S*)-hydrobenzoin (**3b**), (3*S*,4*R*)-2,3-dihydroxy-4-phenylbutan-2-one (**4a**), (3*R*,4*S*)-2,3-dihydroxy-4-phenylbutan-2-one (**4b**), methyl (2*S*,3*R*)-2,3-dihydroxy-3-phenylpropionate (**5a**), methyl (2*R*,3*S*)-2,3-dihydroxy-3-phenylpropionate (**5b**), diethyl D-tartrate (**6a**), and diethyl L-tartrate (**6b**). The diol concentration was 5 mM, and the concentrations of the indicators and the host are listed in Table S1 in the Supporting Information. (B) Absorption spectra of (*R*,*R*)-**2**−ML with diol **5a** (blue ●), racemic **5a/5b** mixture (magenta ◆), and diol **5b** (red ▲). (C) Absorption spectra of (*R*,*R*)-**2**−PV with diol **5a** (blue ●), racemic **5a/5b** mixture (magenta ◆), and diol **5b** (red ▲). All of the studies were carried out in 10 mM *p*-toluenesulfonic acid/Hunig's base buffer (pH 7.4) in 100% MeOH at 25 °C.

Table 2. Screening Plate Results: ΔAbs between the Enantiomeric Guest Diols (5 mM) Using Host (*S*,*S*)-**1** with the Indicators, Calculated at the Given Wavelengths

		guest diol			
indicator	λ (nm)	**3**	**4**	**5**	**6**
A	450	0.007	0.035	0.024	0.007
AC	550	0.026	0.025	0.008	0.014
BPG	570	0.233	0.054	0.119	0.038
ML	380	0.145	0.061	0.102	0.043
PG	520	0.031	0.031	0.030	0.007
PV	520	0.107	0.080	0.072	0.057

Table 3. Screening Plate Results: ΔAbs between the Enantiomeric Guest Diols (5 mM) Using Host (*R*,*R*)-**2** with the Indicators, Calculated at the Given Wavelengths

		guest diol			
indicator	λ (nm)	**3**	**4**	**5**	**6**
A	450	0.176	0.092	0.158	0.043
AC	550	0.161	0.042	0.107	0.002
BPG	570	0.251	0.046	0.156	0.018
ML	380	0.357	0.208	0.352	0.126
PG	520	0.005	0.010	0.031	0.112
PV	520	0.216	0.166	0.232	0.139

nation of all the diols, showing ΔAbs > 0.05 (Tables 2 and 3). The **1**−BPG pair also showed good discrimination between the enantiomers of **3** and **5**, with ΔAbs values of 0.23 and 0.16, respectively, but the discrimination was poor for the enantiomers of diols **4** and **6**. This process allowed us to rapidly identify the best host−indicator combination for the enantiodiscrimination of each chiral diol, eliminating the need to carry out UV−vis titrations of each chiral diol with each host and indicator combination. The total number of UV−vis titrations circumvented is 96, as there are eight diols, two hosts, and six indicators.

Principal Component Analysis. Our first goal was to demonstrate chemo- and enantioselective discrimination of our analytes. As identified by the screening, our assay utilized three

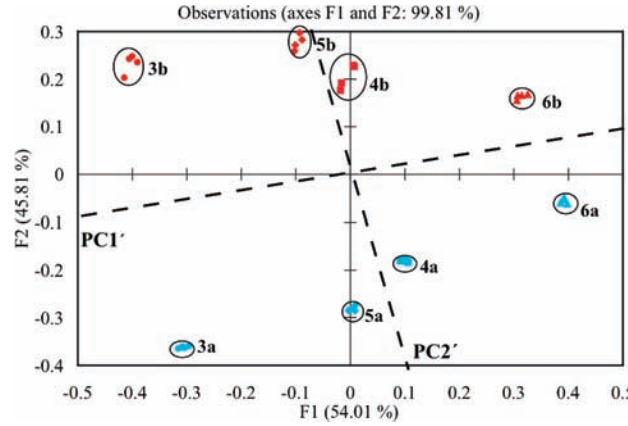


Figure 5. PCA plot for the enantiomers of the four diols analyzed with the host−indicator combinations (*S*,*S*)-**1**−PV, (*R*,*R*)-**2**−ML, and (*S*,*S*)-**2**−PV. The diols are labeled on the PCA plot. All of the studies were carried out in 10 mM *p*-toluenesulfonic acid/Hunig's base buffer (pH 7.4) in 100% MeOH at 25 °C.

host−indicator combinations: (*S*,*S*)-**1**−PV, (*R*,*R*)-**2**−ML, and (*S*,*S*)-**2**−PV. Each diol at 5 mM concentration was treated with these three host−indicator combinations. The experiment was repeated four times to ensure reproducibility. A full spectrum of each sample was recorded, and the data were analyzed at nine different wavelengths: 496, 500, and 516 nm for (*S*,*S*)-**1**−PV, 362, 366, and 374 nm for (*R*,*R*)-**2**−ML, and 496, 500, and 516 nm for (*S*,*S*)-**2**−PV (Table S2 in the Supporting Information). These wavelengths were chosen on the basis of the largest change in absorbance. The XLSTAT computer program was used to generate a PCA plot, which showed excellent discrimination for all of the diols and their enantiomers (Figure 5). Tight clustering of identical samples and good spatial resolution for all of the analytes was achieved. A slight rotation of the PCA axes showed chemoselectivity separated along PC1′ and chirality correlated along PC2′. Guests with *S* stereochemistry at the stereocenter near the phenyl group (or ester group of **6**) have positive scores along PC2′, while those with *R*

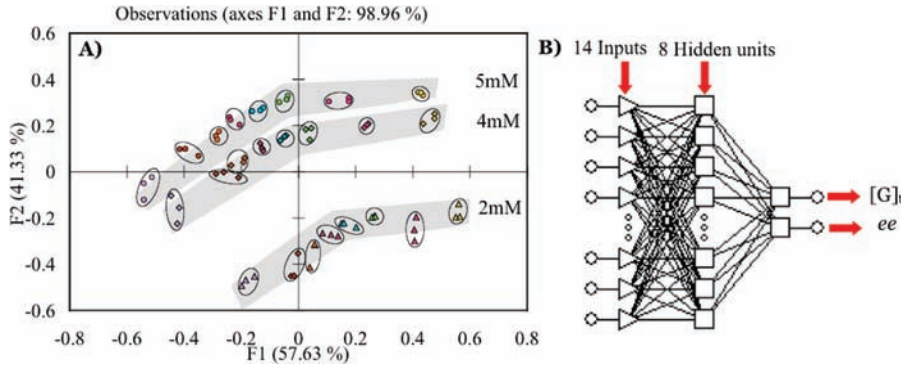


Figure 6. (A) PCA of diol **5** *ee* titration at three different concentrations [2 mM (▲), 4 mM (◆), and 5 mM (●)] for various *ee*'s [−1 (yellow), −0.6 (pink), −0.2 (green), 0 (blue), 0.2 (red), 0.4 (orange), 0.6 (brown), and 1 (purple), where 1 is 100% **5a** and −1 is 100% **5b**]. (B) Multilayer perceptron (MLP) artificial neural network generated for the determination of *ee* and analyte total concentration ($[G]_t$). All studies were carried out in 10 mM *p*-toluenesulfonic acid/Hunig's base buffer (pH 7.4) in 100% MeOH at 25 °C.

stereochemistry have negative scores. Furthermore, the extent of enantioselectivity of the sensor array [(*S*,*S*)-**1**−PV, (*R*,*R*)-**2**−ML, and (*S*,*S*)-**2**−PV] to the diols is also graphed on the PCA plot. As shown, the enantiomers of diol **3** are farther apart than those of the other diols with respect to PC2′, meaning that the array best discriminates the enantiomers of diol **3**. The enantiomers of diol **6**, however, are closer together with respect to PC2′ and are not as well discriminated by the array. This result is supported by the screening-plate analysis, where the observed ΔAbs between diol enantiomers using the (*S*,*S*)-**1**−PV, (*R*,*R*)-**2**−ML, and (*S*,*S*)-**2**−PV combinations was larger for diol **3** than for the other diols. Therefore, the PCA plot also graphically displays how well each diol is enantiomerically discriminated by the array.

Having found excellent chemo- and enantioselectivity, the next goal was to explore the ability to simultaneously determine $[G]_t$ and *ee*. For this purpose, we arbitrarily chose to use diol **5** with the three different host−indicator combinations. The earlier screening experiment discussed above for diol **5** determined that the best enantiomeric discrimination between **5a** and **5b** was achieved by host−indicator combinations **1**−BPG (ΔAbs = 0.119), **2**−ML (ΔAbs = 0.352), and **2**−PV (ΔAbs = 0.232). Thus, *ee* titrations of diol-**5** with only these three combinations were carried out at three different concentrations (2, 4, and 5 mM) and eight different *ee* values (−1.0, −0.6, −0.2, 0, 0.2, 0.4, 0.6, and 1.0, where 1.0 is 100% **5a** and −1.0 is 100% **5b**). Each *ee* titration was carried out at one concentration and repeated three times. The data set consisting of wavelengths 570, 572, 574, 576, and 578 nm for (*S*,*S*)-**1**−BPG; 512, 514, 516, and 518 nm for (*S*,*S*)-**2**−PV; and 376, 378, 380, 382, and 384 nm for (*R*,*R*)-**2**−ML was analyzed using PCA (Table S3 in the Supporting Information). Good spatial resolution was obtained in the PCA plot, which showed clustering of identical samples and spatial resolution of concentration and *ee* (Figure 6A): the data sets with varying *ee* values for a given concentration were clustered together in smooth curves, and the *ee* values ranged within the stripe left to right from +1 to −1, respectively (see the legend of Figure 6A). In this analysis, we did not have to use an achiral host to determine the concentrations of analytes, in contrast to our previous studies. Using data from different host−indicator combinations enabled us to eliminate the achiral host from our sensor array, thus simplifying the analysis.[23]

Artificial Neural Network Analysis. The last goal was to analyze unknown samples for concentration and *ee*. A neural network was generated using the same data set that was used to generate the PCA plot for diol **5** as the ANN training set (Figure 6B). The repetitive data was not included in the training set, as ANN analysis does not work well with redundant values. Statistica Neural Network software was used to develop the ANN. It has an embedded intelligent problem solver (IPS) function that automatically generates several networks that are suitable for the designated problem at hand. The input layer contains the absorbance of each *ee* titration, giving a total of 14 absorbance values. The outputs are $[G]_t$ and % **5a**. ANN analysis does not work well with negative values, so % **5a** was used instead of *ee*. A multilayered perceptron (MLP) network with 14 inputs, eight processing units in the hidden layer, and two outputs was selected for our analysis. The network was trained by back-propagation algorithms, which minimize the discrepancy between the input and the outputs. True unknown samples, prepared completely independently of the training set, were then treated with the same sensor array, and their $[G]_t$ and % **5a** values were predicted by the network (Table S4 in the Supporting Information). The % **5a** values predicted by the network were converted to *ee* and are listed in Table 4. The average absolute errors for $[G]_t$ and *ee* were ±0.08 mM and 6.72%, respectively. The *ee* error was large because of a single outlier having an *ee* error of 19.38%. When the outlier was excluded, the average absolute error in *ee* dropped to 3.57%. The *ee* error could be further reduced by using a larger training set, as shown previously by our group.[30]

Table 4. ANN Analysis of Unknown Solutions

$[G]_t$ (mM)			% *ee*		
exptl	ANN	error	exptl	ANN	error
3.00	3.07	0.07	50.00	69.38	19.38
3.00	3.04	0.04	30.00	33.10	3.10
3.00	3.05	0.05	10.00	8.47	1.53
3.00	3.04	0.04	−30.00	−20.64	9.36
3.00	2.79	0.21	−50.00	−50.27	0.27

Overall, the time required for this entire analysis, including screening, training, and analysis by ANN and PCA, required ∼24 h of work. The method is very rapid once an ANN network has been developed by using a training set for a particular analyte. The concentration and *ee* values of 96 unknown samples can be determined in 40 min, including 30 min for the pipetting system to load a 96-well plate with the host, indicator, and unknown solutions and 10 min for the 96-well plate reader to record the absorbances of the 96 unknown samples at five different wavelengths needed for the ANN analysis. The absorbance data of unknown samples can be channeled to the developed ANN, which in real time gives the concentrations

and *ee*'s of the unknown samples. In effect, this assay is a powerful tool that we are now implementing in a high-throughput study to determine the identity, $[G]_t$, and *ee* of unknown samples of chiral vicinal diols.

Conclusion

We have described a technique that fingerprints chemical identity, concentration, and chirality of chiral vicinal diols. Pattern recognition protocols were used to analyze the data. Excellent fingerprinting in which the diols were chemoselectively and enantioselectively separated was obtained. This was achieved by using a diverse sensor array made up of different host and indicator combinations. We have also shown that a highly enantioselective array that does not require the presence of an achiral boronic acid host can be used to calculate concentrations of unknown samples. A family of chiral hosts and indicator combinations with an ANN can provide both concentration and *ee* values of unknown samples with high accuracy. In addition, a new enantioselective boronic acid host **2** showing excellent enantioselectivity for several chiral vicinal diols has been designed and synthesized in one step. To study the enantioselective behavior of our boronic acid host, we solved an X-ray structure of **7**, the complex of host (*R*,*R*)-**2** with (*S*,*S*)-hydrobenzoin (**3b**). Studies of the solution-phase behavior of the complex by ^{11}B NMR lent some insight into the enantioselective behavior of host **2** and suggested that the hybridization of the boronic acid species depends on both the solvent and the substrate as well as on the structure of the boronic acid receptor.

Acknowledgment. We gratefully acknowledge support from the NIH (5 RO1 GM077437) and the Welch Foundation (F-1151) for funding of this work.

Supporting Information Available: Spectroscopic data for the host (*S*,*S*)-**2** with the indicators; UV−vis data for PCA and ANN; crystallographic data for complex **7** (CIF); and ^{11}B NMR spectra of host (*R*,*R*)-**2** and complex **7**. This material is available free of charge via the Internet at http://pubs.acs.org.

JA904545D