J|A|C|S ARTICLES

Published on Web 07/30/2009

The Use of Differential Receptors to Pattern Peptide Phosphorylation

Tianzhi Zhang,[†] Nicola Y. Edwards,[‡] Marco Bonizzoni,[†] and Eric V. Anslyn*[,†]

Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78712, and Department of Chemistry, Pennsylvania State Worthington Scranton, Dunmore, Pennsylvania 18512

Received May 22, 2009; E-mail: anslyn@austin.utexas.edu

Abstract: An array sensing scheme for the differentiation of small peptides and their phosphorylated analogues is introduced. The technique involves a series of receptors created by appending random peptides to a C_{3v} symmetric scaffold that binds phosphomonoesters. Five specific peptide sequences were selected through a screening technique. In addition to cross reactivity being created by the peptides in the receptors, three metal ions and three pH indicators are used to create a suite of 45 indicator displacement assays. The colorimetric data from the 45 sensing ensembles is collected in a 96-well plate reader, and linear discriminant analysis gives patterns resulting in 100% classification of the peptides. The approach demonstrates a generalizable principle to create pattern-based recognition protocols for complex analytes.

1. Introduction

Post-translational modification has profound effects on the structure, function, activity, and molecular recognition properties of proteins. A broad repertoire of cellular functions including metabolic control, gene expression, signal transduction, cell division and differentiation, as well as apoptosis, are influenced by these modifications.[1] Protein phosphorylation is the most common post-translational modification. It is estimated that at any one time 40% of all proteins in eukaryotic cells are phosphorylated.[2] Switching phosphorylation and dephosphorylation on and off in the cell by using inhibitors and activators of kinase and phosphatase enzymes is a promising course of treatment in several diseases.[3] Hence, identifying sites of phosphorylation on proteins is critical to the understanding of key signaling pathways which are involved in the onset and progression of certain disorders.[4]

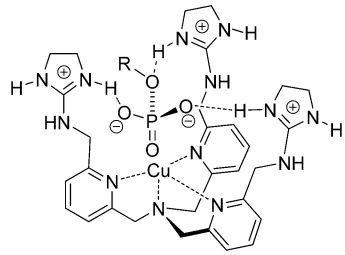
Simple and efficient detection schemes could serve as tools for monitoring post-translational modifications and as screens for kinase/phosphatase activators and inhibitors. Efforts in this arena have largely focused on the use of fluorescent probes appended to peptides or synthetic analogues that are kinase substrates. The probes signal phosphorylation of the substrate.[5] The range of phosphorylated species that could be detected would be significantly broadened with the use of probe-free substrates by utilizing molecular receptors. For example, Hamachi and co-workers have designed elegant sensors utilizing dinuclear zinc complexes with fluorophores whose signal is modulated upon binding phosphorylated peptides and other species.[6] Their efforts have resulted in sensors that discriminate phosphorylated from nonphosphorylated peptides and do so in a sequence-dependent fashion. Analogously, a highly selective hybrid biosensor has been developed by the same group for monitoring a kinase-catalyzed phosphorylation.[7]

We have previously reported a C_{3v}-symmetric receptor [Cu^{II}(**1**)] that was found to be highly selective in the binding of tetrahedral oxoanions, such as inorganic phosphate, over other anions.[8] The binding of a Cu^{II} ion to **1** creates a relatively rigid and preorganized cavity that is complementary to tetrahedral oxoanions. Modeling predicts that one oxygen atom of phosphate binds to the metal ion; this is accompanied by individual hydrogen-bonding contacts of the remaining three oxygen atoms of phosphate with the three guanidinium groups of the receptor. Our previous studies showed good affinity of [Cu^{II}(**1**)] for phosphate and arsenate. We further found that the receptor binds monophosphoesters, such as those of serine, threonine, and tyrosine, with only a slightly decreased affinity relative to phosphate.[9]

[†] University of Texas.
[‡] Penn State Worthington Scranton.

(1) Doyle, H. A.; Mamula, M. J. *Trends Immunol.* **2001**, *22*, 443–449. Krishna, R. G.; Wold, F. *Proteins* **1998**, 121–206.
(2) Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M. *Nat. Biotechnol.* **2002**, *20*, 301–305.
(3) Burke, R. E. *Pharmacol. Ther.* **2007**, *114*, 261–277. Churchill, E.; Budas, G.; Vallentin, A.; Koyanagi, T.; Mochly-Rosen, D. *Ann. Rev. Pharmacol. Toxicol.* **2008**, *48*, 569–599. Lee, M. R.; Duan, W.; Tan, S.-L. *Expert Opin. Ther. Targets* **2008**, *12*, 535–552. Tamura, T.; Koch, A. In *Kinase Targets and Inhibitors in Inflammation*; Asadullah, K., Gaestel, M., Eds.; Bentham Science Publishers: Bussum, The Netherlands, 2007; pp 87−114.
(4) Anderson, J. C.; Peck, S. C. *Plant J.* **2008**, *55*, 881–885. Riechers, A.; Schmidt, F.; Stadlbauer, S.; Koenig, B. *Bioconjugate Chem.* **2009**, *20*, 804–807. Roche, K. W.; Huganir, R. L. *Current Protocols in Neuroscience*; Crawley, J. N., Ed.; Wiley: New York, 2001; Chapter 5, Unit 5-14.
(5) Chen, C.-A.; Yeh, R.-H.; Lawrence, D. S. *J. Am. Chem. Soc.* **2002**, *124*, 3840–3841. McIlroy, B. K.; Walters, J. D.; Johnson, J. D. *Anal. Biochem.* **1991**, *195*, 148–52. Yeh, R.-H.; Yan, X.; Cammer, M.; Bresnick, A. R.; Lawrence, D. S. *J. Biol. Chem.* **2002**, *277*, 11527–11532.
(6) Ojida, A.; Mito-oka, Y.; Inoue, M.-a.; Hamachi, I. *J. Am. Chem. Soc.* **2002**, *124*, 6256–6258.
(7) Anai, T.; Nakata, E.; Koshi, Y.; Ojida, A.; Hamachi, I. *J. Am. Chem. Soc.* **2007**, *129*, 6232–6239.
(8) Tobey, S. L.; Jones, B. D.; Anslyn, E. V. *J. Am. Chem. Soc.* **2003**, *125*, 4026–4027.
(9) Zhang, T.; Anslyn, E. V. *Tetrahedron* **2004**, *60*, 11117–11124.

10.1021/ja9041675 CCC: $40.75



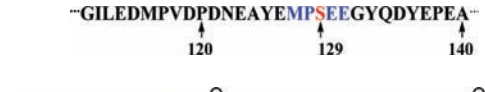
[Cu^{II}(**1**)(OPO_2OR)]

Although detection schemes for specific kinase activity could be developed using molecular receptors that selectively bind to and signal the corresponding phosphorylated peptides/proteins, one would need a different receptor for each phosphorylated substrate. Alternatively, the use of arrays of receptors could fingerprint phosphorylation events and thereby pattern the peptide/protein substrates and their phosphorylated products. Our group,[10−12] and others,[13] have extensively explored pattern-based array sensing, often of biologically relevant analytes. This area of research, coined *differential sensing*,[14] has emerged as a powerful tool for detecting chemically diverse analytes, for discriminating subtly different targets, and for classifying mixtures.[15] Discrimination is achieved through the use of an array of cross-reactive sensor elements that interact with analytes in a differential fashion, to produce a composite signal (i.e., a fingerprint) for an analyte or mixtures of interest.

Our approach to differential sensing employs recognition moieties generically targeted toward the recurring binding motif present in the analytes of interest, combined with interacting pendant arms bearing amino acid libraries which impart the necessary cross reactivity.[16] These combinatorially derived receptors are organized into an array sensing platform, and the binding of analytes is assessed optically using colorimetric/fluorescent indicator uptake or displacement protocols. Patterns within the optical data sets are identified using pattern recognition algorithms.

As a proof of principle for detecting protein phosphorylation, we first targeted smaller peptides and their phosphorylated versions. As a choice of peptides, we examined one involved in Parkinson's disease (PD). This disease is a motor system disorder which results from the loss of dopamine-producing brain cells.[17] A phosphorylated derivative of filamentous α-synuclein has been identified as the main component of Lewy bodies, protein aggregates found in the brains of people who suffered from PD. The dominant pathological modification of α-synuclein has been shown to be phosphorylation at Ser-129 (Scheme 1).[18]

Scheme 1. Partial Amino Acid Sequence of Filamentous α-Synuclein, Including the Ser129 Phosphorylation Site Believed to Be Responsible for Anomalous Aggregation and Precipitation in Lewy-Body Neurological Diseases



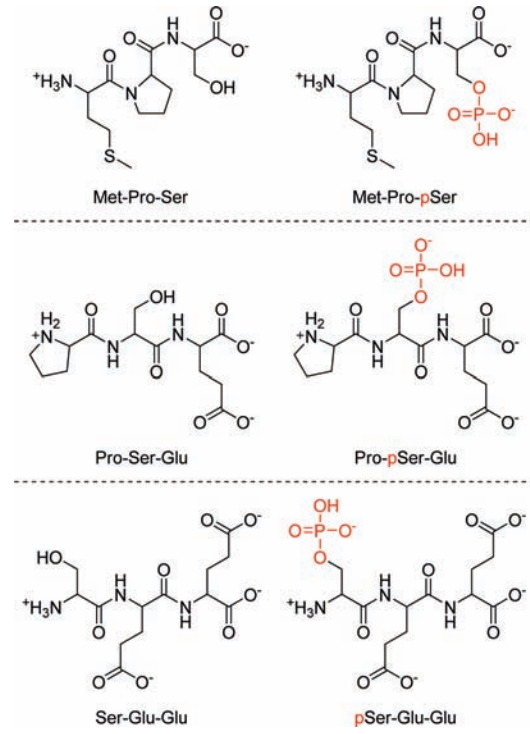


Figure 1. Target tripeptides and their phosphorylated serine derivatives considered in this study, in the prevalent protonation states at physiological pH.

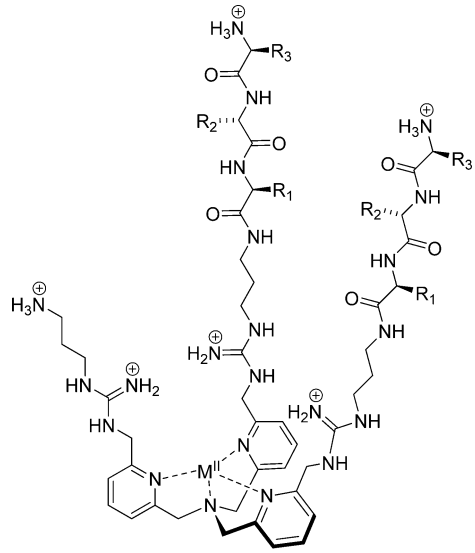
Synthetic peptides derived from the amino acid sequence surrounding Ser-129 of filamentous α-synuclein (specifically the Met-Pro-Ser-Glu-Glu motif) were used as analytes to assess the discriminatory properties of our differential sensor protocols. The goal was to determine if our approach to sensing could discriminate tripeptides that flank and contain the site of phosphorylation (Met-Pro-Ser = MPS, Pro-Ser-Glu = PSE and Ser-Glu-Glu = SEE) from their phosphorylated serine derivatives (Met-Pro-pSer = MPS_P, Pro-pSer-Glu = MS_PE and pSer-Glu-Glu = S_PEE) (Figure 1).

2. Results and Discussion

2.1. Design Criteria. Synthetic receptor [Cu^{II}(**1**)] was the starting point for our differential receptor library design. We have previously taken other molecular receptors developed for specific functional groups, such as triphosphates, amino acids, and sugars, and by attachment of random peptides to these scaffolds, we have been able to fingerprint nucleosides,[11] short peptides,[10] and glycoproteins,[12] respectively. Important for this study is the fact that we have also previously shown that a

(10) Folmer-Andersen, J. F.; Kitamura, M.; Anslyn, E. V. *J. Am. Chem. Soc.* **2006**, *128*, 5652–5653.
(11) McCleskey, S. C.; Griffin, M. J.; Schneider, S. E.; McDevitt, J. T.; Anslyn, E. V. *J. Am. Chem. Soc.* **2003**, *125*, 1114–1115.
(12) Wright, A. T.; Griffin, M. J.; Zhong, Z.; McCleskey, S. C.; Anslyn, E. V.; McDevitt, J. T. *Angew. Chem., Int. Ed.* **2005**, *44*, 6375–6378.
(13) Margulies, D.; Hamilton, A. D. *Angew. Chem., Int. Ed.* **2009**, *48*, 1771–1774. Palacios, M. A.; Wang, Z.; Montes, V. A.; Zyryanov, G. V.; Anzenbacher, P., Jr. *J. Am. Chem. Soc.* **2008**, *130*, 10307–10314. Sandanaraj, B. S.; Demont, R.; Thayumanavan, S. *J. Am. Chem. Soc.* **2007**, *129*, 3506–3507.
(14) Lavigne, J. J.; Anslyn, E. V. *Angew. Chem., Int. Ed.* **2001**, *40*, 3118–3130. Wright, A. T.; Anslyn, E. V. *Chem. Soc. Rev.* **2006**, *35*, 14–28.
(15) Goodey, A.; Lavigne, J. J.; Savoy, S. M.; Rodriguez, M. D.; Curey, T.; Tsao, A.; Simmons, G.; Wright, J.; Yoo, S.-J.; Sohn, Y.; Anslyn, E. V.; Shear, J. B.; Neikirk, D. P.; McDevitt, J. T. *J. Am. Chem. Soc.* **2001**, *123*, 2559–2570.
(16) Edwards, N. Y.; Sager, T. W.; McDevitt, J. T.; Anslyn, E. V. *J. Am. Chem. Soc.* **2007**, *129*, 13575–13583.
(17) Cookson, M. R. *Annu. Rev. Biochem.* **2005**, *74*, 29–52. Marsden, C. D. *Lancet* **1990**, *335*, 948–952. Olanow, C. W.; Tatton, W. G. *Annu. Rev. Neurosci.* **1999**, *22*, 123–144.
(18) Anderson, J. P.; et al. *J. Biol. Chem.* **2006**, *281*, 29739–29752. Fink, A. L. *Acc. Chem. Res.* **2006**, *39*, 628–634. Sugeno, N.; Takeda, A.; Hasegawa, T.; Kobayashi, M.; Kikuchi, A.; Mori, F.; Wakabayashi, K.; Itoyama, Y. *J. Biol. Chem.* **2008**, *283*, 23179–23188.

prescreening of the random peptides significantly improves the cross reactivity, and hence the patterning, obtained from the array technique.[19] Anticipating that similar strategies could fingerprint phosphopeptides, we designed the family of combinatorial receptors [M^{II}(**2**)].



[M^{II}(**2**)]

Structure **2** possesses the same amine-tris(pyridine) ligand and three guanidinium groups as does **1**, but in addition, two random tripeptide chains have been attached to the guanidinium groups. The remaining guanidinium moiety is connected to a simple primary amine, a residual linker from the solid-phase synthesis procedure (see below). Since we have found previously that phosphomonoesters bind to [Cu^{II}(**1**)] with good affinity,[9] we envisioned a comparable binding to the similar motif found in [M^{II}(**2**)]. Upon binding of a phosphoserine peptide to [M^{II}(**2**)], the two random peptidic arms are postulated to present different electrostatic, hydrophobic, and hydrogen-bonding interactions with the side chains of the target tripeptides (Figure 1). As described below, screening of the peptide-containing receptor library winnows down the receptors to those that successfully differentiate the targets.

Furthermore, by varying the bound metal ion we can easily introduce additional differential binding toward the peptide targets. Although our previous studies all centered on the use of Cu^{II}, it is well-known that the tris-pyridine ligand used in **1** and **2** binds several transition metals.[20] Different metals inherently impart different affinities for phosphate and, in some cases, may also change the coordination geometry of the ligand, thereby imparting different structures to the [M(**2**)] hosts.

The detection method to be employed must operate at low millimolar levels. At such concentration, the absorption from the metal-centered d−d transition in the visible region is very small. To obtain accurate and reproducible data, we envisioned the use of indicator displacement assays (IDAs)[21] to afford a much more intense chromophoric reporter. The modularity of an IDA[22,23] allows us to employ several different indicators, further increasing the diversity of the elements in our array that will carry out the fingerprinting of the six target tripeptides.

In summary, the design that went into our array for differential patterning of phosphopeptides and corresponding nonphosphorylated peptides incorporates three levels of diversity: different transition metal ions, different indicators, and random peptide sequences on two of the receptor arms.

2.2. Screening of Indicators and Metal Ions. Dozens of indicators and metal ions could be considered for inclusion in the present study, and although in principle having a large number of components in the system would ensure great diversity, this approach would be impractical due to the sheer number of combinations to be tested. We therefore set out to identify and screen a limited number of candidate metal ions and indicators in order to find the most promising combinations for our purpose.

For the sake of simplicity, the indicator and metal screening phase of the study was carried out using only the core structure of **2**. Structure **1**, previously reported by our group as having high affinity for the phosphate anion, was used as a model of the metal-and-anion binding core of compounds **2**. Similarly, simple phosphoserine was used as representative of our ultimate phosphorylated tripeptide targets.

We selected a number of indicators containing metal coordination sites (e.g., catechol and carboxylate moieties) to be included in the screening pool (Figure 2). Furthermore, divalent and trivalent first-row transition metal ions were considered as metal ion candidates, whose differences in size, charge, and preference of coordination geometry are expected to create a differential response in the array. Specifically, Cr^{III}, Mn^{II}, Fe^{III}, Fe^{II}, Co^{II}, Ni^{II}, Cu^{II}, and Cd^{II} chlorides were evaluated.

The screening was conducted in water (Tris buffer, pH 7.4) in 96-well microtiter plates. All combinations of model compound **1** with the eight metal ions and nine indicators were prepared, giving rise to $8 \times 9 = 72$ trials. Two sets of such solutions were prepared on separate 96-well plates, with and without phosphoserine, respectively. Two indicators (methyl red and mordant orange 1) and Cd(II) were discarded outright because little to no spectral modulation was observed in all combinations.

For all other indicator and metal ion combinations, the highest difference in absorbance between the solutions with and without phosphoserine was calculated. Of course, the wavelength at which the reported maximum change was observed is different for each indicator and metal ion combination. However, the specific position of the maximum of change was not considered informative for the final purpose of pattern recognition: instead, the amount of change in absorbance was deemed to contain more useful dif-

(19) Wright, A. T.; Edwards, N. Y.; Anslyn, E. V.; McDevitt, J. T. *Angew. Chem., Int. Ed.* **2007**, *46*, 8212–8215.

(20) Adams, H.; Bailey, N. A.; Fenton, D. E.; He, Q.-Y. *J. Chem. Soc., Dalton Trans.* **1997**, 1533–1539. Anderegg, G.; Hubmann, E.; Podder, N. G.; Wenk, F. *Helv. Chim. Acta* **1977**, *60*, 123–140. Blackman, A. G. *Eur. J. Inorg. Chem.* **2008**, 2633–2647. Cheyne, S. E.; McClintock, L. F.; Blackman, A. G. *Inorg. Chem.* **2006**, *45*, 2610–2618. Chuang, C.-L.; dos Santos, O.; Xu, X.; Canary, J. W. *Inorg. Chem.* **1997**, *36*, 1967–1972. Dai, Z.; Canary, J. W. *New J. Chem.* **2007**, *31*, 1708–1718. Di Bernardo, P.; Melchior, A.; Portanova, R.; Tolazzi, M.; Zanonato, P. L. *Coord. Chem. Rev.* **2008**, *252*, 1270–1285. Neubrand, A.; Thaler, F.; Koerner, M.; Zahl, A.; Hubbard, C. D.; van Eldik, R. *J. Chem. Soc., Dalton Trans.* **2002**, 957–961. Pettinari, C.; Marchetti, F.; Drozdov, A. In *Comprehensive Coordination Chemistry II: From Biology to Nanotechnology*; McCleverty, J. A., Meyer, T. J., Eds.; Elsevier Pergamon: Amsterdam, Boston, 2004; Vol. 1, pp 211−251.

(21) Nguyen, B. T.; Anslyn, E. V. *Coord. Chem. Rev.* **2006**, *250*, 3118–3127.

(22) Leung, D.; Folmer-Andersen, J. F.; Lynch, V. M.; Anslyn, E. V. *J. Am. Chem. Soc.* **2008**, *130*, 12318–12327.

(23) Shabbir, S. H.; Regan, C. J.; Anslyn, E. V. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10487–10492.

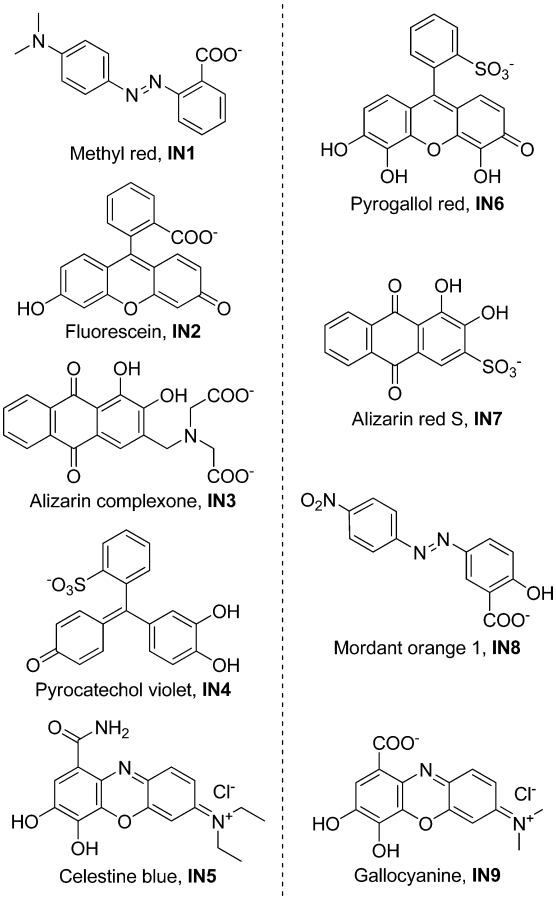


Figure 2. Indicators subjected to the screening process.

ferential information. Consequently, for each combination of receptor **1**, indicator, and metal ion, the reported signal (Figure 3) was calculated as the maximum value in a difference spectrum obtained by subtracting the absorbance spectrum measured in the presence of plain serine from the one acquired with phosphoserine: signal = $\max(\mathrm{Abs}_{\mathrm{pSer}}(\lambda) - \mathrm{Abs}_{\mathrm{Ser}}(\lambda))$. The purpose of this screening was not to maximize affinity or selectivity, but on the contrary, to maximize cross reactivity and *differential* response. Along these lines, inspection of Figure 3 shows that gallocyanine's absorbance (**IN9**) changes significantly between Cu^{II} and Ni^{II}, while pyrocatechol violet (**IN4**) shows differences for Ni^{II} compared to Cu^{II} and Co^{II}, and last, celestine blue (**IN5**) shows significant differences for all three metal ions Ni^{II}, Cu^{II}, and Co^{II}. Metal and indicator combinations that consistently showed negative absorbance changes throughout were not considered further, because this is not consistent with a simple indicator displacement model, and is probably indicative of more complex behavior. Even though such a response could prove useful in the construction of a differential array, we decided to restrict the scope to combinations providing results consistent with the IDA mechanism, in an effort to keep the system as simple as possible, without compromising its discriminatory power.

In summary, the indicators pyrocatechol violet (**IN4**), celestine blue (**IN5**), and gallocyanine (**IN9**), and the metal ions Ni^{II}, Co^{II}, and Cu^{II} were selected because, in combinations with **1**, they exhibited the largest differences between their respective UV/vis spectra when exposed to phosphoserine.

2.3. Synthesis of the Resin-Bound Tripodal Scaffold. After identification of the most promising indicators and metal ions (see section 2.2 above), the first stage of our search for effective receptors involved the synthesis of a version of receptor **1** amenable to attachment to a resin for solid-phase peptide coupling, resulting in the design of **3**. Attachment of the receptors to a resin was required in order to conduct combinatorial syntheses allowing access to a library of peptides as denoted in the general structure of **2**. The preparation of the library was achieved by standard split-and-pool combinatorial techniques.[24] Moreover, we planned to conduct a prescreening of the obtained library while it was still attached to the solid support (see below), in order to identify the most promising candidates out of the generated combinatorial pool, prior to the use of the library in an array. This 2-fold goal required that the starting compound be desymmetrized, in order to use a single arm bearing a free amine moiety as an attachment point to the resin. This was achieved through opportune use of Boc- or Fmoc-protected ω-aminothioureas as guanidylating agents.

Our synthetic efforts built upon the previously reported structure **4** (Figure 4).[8,9] Monoprotection of **4** with Boc-ON afforded compound **5**, which was then reacted with two equivalents of the protected thiourea **6** in the presence of the coupling agent *N*-ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide (EDCI) to produce intermediate **7**. Removal of the Boc protecting group from **7** was effected with TFA and afforded compound **8**. Reaction of **8** with thiourea **9** under conditions similar to those reported above generated intermediate **10**, which finally afforded product **3** through removal of the remaining Boc group with TFA.

2.4. Combinatorial Library of Peptidic Receptors. The second phase of the synthesis was the creation of the desired receptor library. The combinatorial library **2**, consisting of $19^3 = 6859$ receptors, was synthesized using the readily available amine-functionalized Tentagel resin as solid support. In order to load the amine-functionalized scaffold **3** onto this resin, we needed to modify the resin surface groups with a simple linker. Many linker moieties have been developed for resin decoration: however, since the resulting resin-bound library was intended to be used solely in our solid-phase screening process, we did not need to cleave it from the resin at this stage of the synthesis. For this reason, the noncleavable linker α-bromoacetate (BAC) was attached using standard peptide coupling conditions, thus affording resin-attached bromomethyl groups susceptible to nucleophilic attack.[25] A negative result in the ninhydrin test after the BAC coupling confirmed reaction completion and nearly quantitative conversion of the resin amines into bromoacetamides.[26]

The beads were then exposed to a solution containing an excess of structure **3**: loading of **3** onto the resin was attained through alkylation of its terminal primary amine. Subsequently, the Fmoc groups were removed. At this point a ninhydrin test confirmed regeneration of amine termini.

Successively, all the naturally occurring amino acids (except cysteine) were used in every step as building blocks for the peptidic arms (Figure 5) utilizing a split-and-pool paradigm. The resin was subdivided in portions (the *splitting* step) and subjected to standard Fmoc-based peptide coupling conditions

(24) Chan, W. C., White, P. D., Eds. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Oxford University Press: New York, 2000.
(25) See Supporting Information for full experimental details.
(26) Yan, B. *Analytical Methods in Combinatorial Chemistry*; Technomic Publishing Co.: Lancaster, PA, 2000; Gaggini, F.; Porcheddu, A.; Reginato, G.; Rodriquez, M.; Taddei, M. *J. Comb. Chem.* **2004**, *6*, 805–810.

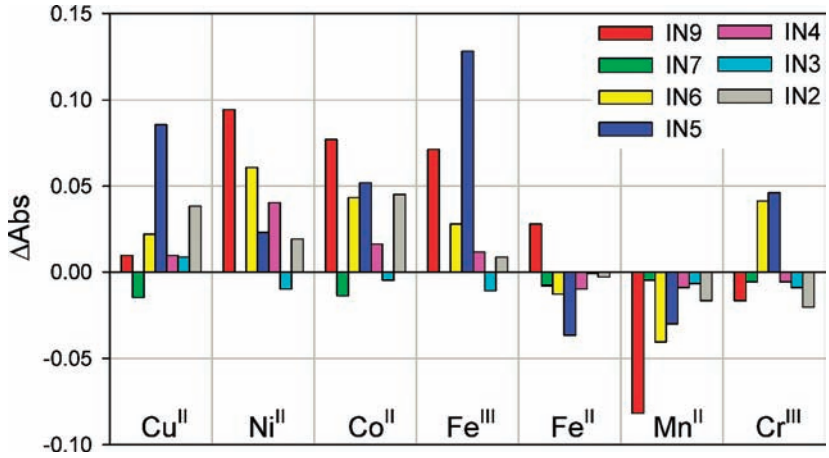


Figure 3. Indicator and metal ion screening: changes in absorbance of a [M(**1**)(**In**)] ensemble in the presence of phosphoserine (pSer). pSer plate: $[M^{n+}] = [In] = [\mathbf{1}] = 3 \times 10^{-5}$ M, [pSer] $= 3 \times 10^{-4}$ M; "blank" plate: $[M^{n+}] = [In] = [\mathbf{1}] = 3 \times 10^{-5}$ M, no pSer. Reported absorbance change = max(Abs[pSer plate] − Abs["blank"]), also see text. All solutions were prepared in water, buffered with Tris (50 mM) to pH 7.4, at 25 °C.

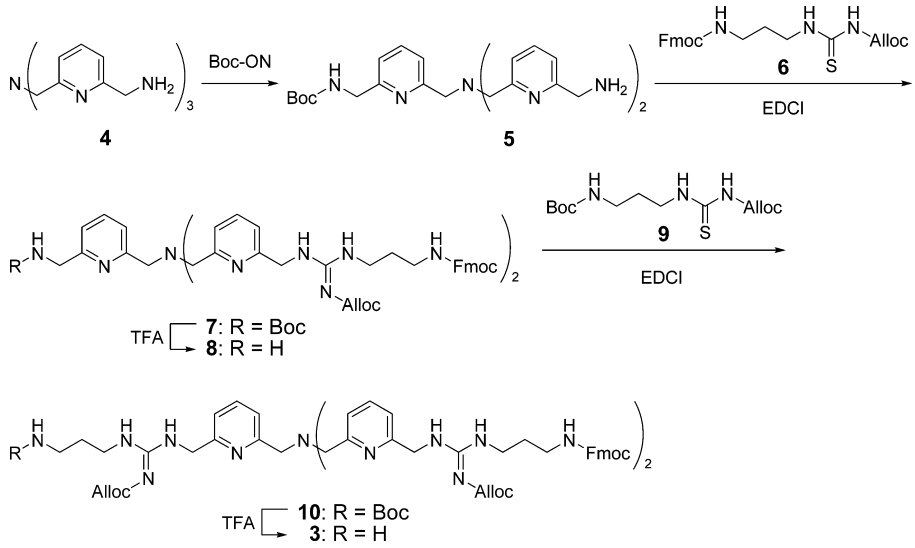


Figure 4. Synthetic route to compound **3**. Detailed procedures and syntheses of compounds **6** and **9** can be found in the Supporting Information.

to attach the first amino acid residues; then all the beads were mixed back together (the *pooling* step), and the process was repeated. A total of three rounds of coupling were performed, and the protected library **11** was obtained on the resin beads. In the final step, Pd^0-catalyzed removal of the Alloc protection from the guanidine groups afforded the desired peptidic receptor library **2** immobilized on the resin.[27]

2.5. Receptor Prescreening. On-bead screening of vast chemical libraries obtained through combinatorial synthesis is a nontrivial problem. A common approach demonstrated by the Still group involves chemically linking a dye to an adequate substrate and treatment of the bead-supported library with dilute solutions of the dye−substrate conjugate. In these conditions, the dye conjugate will be mainly concentrated on beads containing high-affinity receptors, intensely coloring them. Such colored beads can then be picked and the structures of the promising receptors determined.[28]

Building on established expertise in our group instead, we tackled the problem of library screening, and ultimately of anion sensing in the final array, using indicator displacement assays (IDAs).[21] The key requirement is that the selected indicator show different signals when either free in the bulk solution or bound to the host. By far the most relevant advantage of the IDA method to this study is the fact that it does not require synthetic incorporation of the signaling unit in the receptor scaffold, and hence a number of different indicators can be trivially tested in order to optimize the binding properties and the optical signal.[23] This further leads to a modular approach to the construction of an array: several indicators can be used with one receptor, thereby increasing array diversity without requiring additional covalent bond architecture.

We set forth to identify the most promising receptors in library **2**, i.e. the most efficient peptide sequences to be used in the final assay, through screening of the resin-bound receptors. As required by our design (see above), a metal ion was first complexed to the receptors by soaking the resin in a solution of $CuCl_2$ in buffered water to give the metalated library [Cu^{II}(**2**)]. We arbitrarily chose Cu^{II} for the screening since our group had previously established that it binds tightly to the core tetramine ligand and that the resulting complex is capable of binding anions to the vacant apical position.[8] Similarly, out of the three

(27) Gomez-Martinez, P.; Dessolin, M.; Guibe, F.; Albericio, F. *J. Chem. Soc., Perkin Trans. 1* **1999**, 2871–2874.
(28) Still, W. C. *Acc. Chem. Res.* **1996**, *29*, 155–163. Chen, C.-T.; Wagner, H.; Still, W. C. *Science* **1998**, *279*, 851–853.

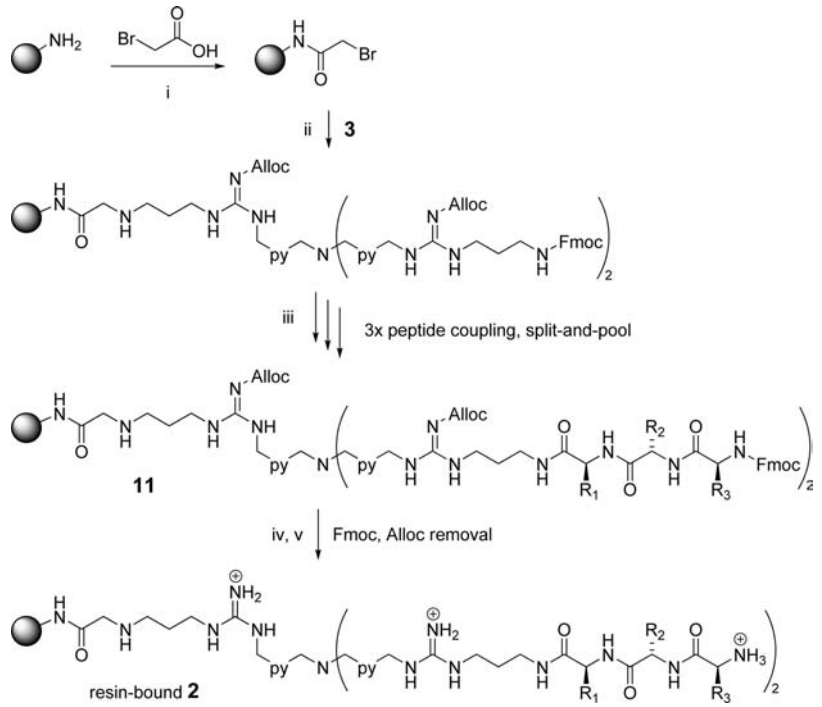


Figure 5. Solid-phase synthesis procedure for the creation of the resin-bound receptor library **2** (py = 2,6-pyridyl linker). (i) diisopropylcarbodiimide, HOBT, *N*-methylmorpholine, dry DMF; (ii) K_2CO_3, dry DMF; (iii) HOBT, TBTU, *N*-methylmorpholine, dry DMF; (iv) 20% piperidine in DMF; (v) $Me_2NH \cdot BH_3$, $Pd^0(PPh_3)_4$ (cat.) in CH_2Cl_2.

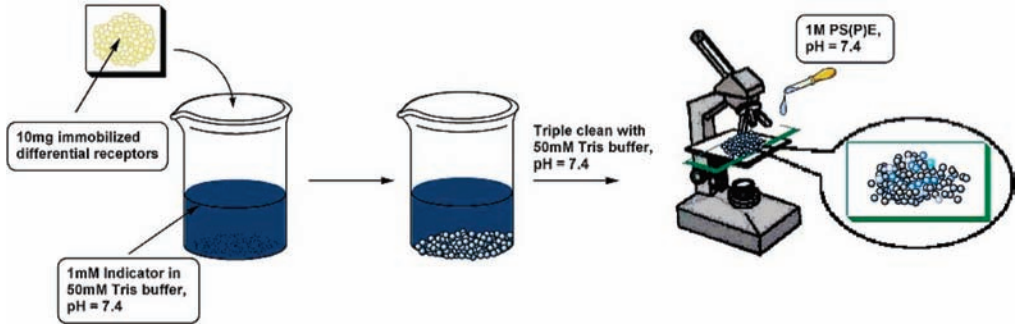


Figure 6. Prescreening procedure: selecting receptors that show good indicator uptake together with effective indicator displacement by phosphorylated tripeptides.

indicators found to be effective for our system (see section 2.2), celestine blue (**IN5**) was arbitrarily chosen to be used in the present step.

We then performed indicator displacement assays with resin-bound receptors [Cu^{II}(**2**)] and celestine blue (**IN5** in Figure 2) (Figure 6). Resin-bound receptors from library [Cu^{II}(**2**)] (10 mg) were incubated in a solution of indicator (1 mM in 50 mM Tris at pH 7.4) for 10 min. Several washing cycles were then performed in order to remove nonspecifically bound indicator. At this stage, a high degree of bead staining indicated high affinity between celestine blue and the corresponding unique resin-bound peptidic arm sequence. Tens of highly stained beads were manually selected and placed in a solution of the phosphorylated Pro-pSer-Glu tripeptide (1 M in 50 mM Tris at pH 7.4). The color of several of these beads was significantly lightened after a 10-min incubation period, indicating displacement of the indicator by Pro-pSer-Glu. Fourteen beads were manually selected and the bound peptidic arms of the receptor sequenced by Edman degradation to reveal their amino acid sequence.

The sequences that were found to be effective are given in Table 1. They display several similarities. All sequences except Met-Ala-Ala (entry 2) have at least one amino acid residue with a polar side chain. All sequences except Lys-Asp-Lys (entry 12) contain amino acids with nonpolar or hydrophobic side chains. Interestingly, amino acid residues with polar side chains are mainly found in positions R_1 and R_2 of all sequences, whereas amino acids with largely hydrophobic or nonpolar side chains are found in position R_3. Proline residues account for five of the R_3 positions.

The five sequences in boldface type in Table 1 were thought to adequately represent the diversity in amino acid nature and position found to be effective in the prescreening experiment, and hence were selected to be appended to the tetraamine core scaffold for the construction of the differential receptors ultimately used in the assay (see below). Hereafter, these sequences will be referred to as $\mathbf{LG_1}$ through $\mathbf{LG_5}$.

2.6. Resynthesis of Selected Receptors. Once appropriate peptidic arm sequences were found, receptors bearing those peptide sequences were selectively resynthesized for use in the

Table 1. Prescreening: Receptors' Peptide Arm Sequencing Results[a]

entry	R_1	R_2	R_3
1 (LG_1)	**Thr**	**Glu**	**Val**
2	Met	Ala	Ala
3 (LG_2)	**Tyr**	**Arg**	**Glu**
4	Asn	Ile	Leu
5	Gln	Thr	Pro
6	Tyr	Val	Pro
7 (LG_3)	**Arg**	**Leu**	**Leu**
8	Asp	Phe	Phe
9	Glu	Ser	Pro
10	Ile	Arg	Pro
11	Phe	Asp	Glu
12 (LG_4)	**Lys**	**Asp**	**Lys**
13	Glu	Asp	Pro
14 (LG_5)	**Tyr**	**Asp**	**Pro**

[a] Peptide sequences obtained through Edman degradation of the tripeptide arms of the selected resin-attached receptors that showed good indicator uptake combined with effective indicator displacement by the phosphorylated tripeptide Pro-pSer-Glu during prescreening.

envisioned differential response array. The synthesis was once again conducted on solid phase for expediency (Figure 5). However, since the resulting products were intended for use in a homogeneous solution assay, and hence had to be cleaved from the solid substrate, we turned to the 2-chlorotrityl chloride resin, which contains an acid labile linker.[29] In this case, compound **3** was directly attached to the resin. Three rounds of peptide coupling followed; here, however, the amino acids used for the construction of the peptide arms were the ones selected from the prescreening experiment, in the order specified by the sequencing study. Finally, the receptors were cleaved from the trityl resin by treatment with TFA.

Five different receptors were prepared (see also Table 1). In all cases HPLC analysis showed single peaks, and mass spectrometry identified the expected molecular ions. The syntheses were remarkably clean, also confirming that, in all likelihood, the same synthetic steps had successfully occurred during the previous combinatorial synthesis on Tentagel beads.

2.7. Array Construction and Fingerprinting. Through the screening procedures outlined in sections 2.2 and 2.5 above, suitable sequences for the receptors' peptidic arms and adequate indicators and transition metal ions were identified for the construction of an array of differential sensors responding to phosphorylated control structures.

Diversity in peptidic receptors (5), metal salts (3), and indicators (3) led to an array of 45 receptor/metal/indicator combinations, expected to produce sufficient differences in absorbance spectra that unique patterns could be obtained for phosphorylated peptides and their nonphosphorylated counterparts (thereby reporting on phosphorylation), and potentially further exploited for identification of the targets.

Data for the fingerprinting assay was again collected using microwell plates. We prepared a 96-well plate for each phosphorylated and nonphosphorylated target tripeptide pair. Each well on the plate contained one of the possible combinations of selected receptor (**LG_1**−**LG_5**), indicator (**IN4**, **IN5**, **IN9**) and metal ion (Co^{II}, Ni^{II}, Cu^{II}), with either the phosphorylated or the nonphosphorylated target tripeptide, for a total of 90 unique combinations. In detail, to each peptidic receptor solution **LG_1**−**LG_5** (3×10^{-5} M), a metal salt (3×10^{-5} M) was added, and the solution was allowed to stand for 5 min; addition of an indicator (3×10^{-5} M) followed. Two sets of such solutions were prepared, and phosphorylated or nonphosphorylated tripeptide substrates (3×10^{-4} M) were then added to the first and second set, respectively. Absorbance spectra were then recorded for the two sets. Nine experimental replicates were obtained for all peptidic substrates for later classification using linear discriminant analysis.

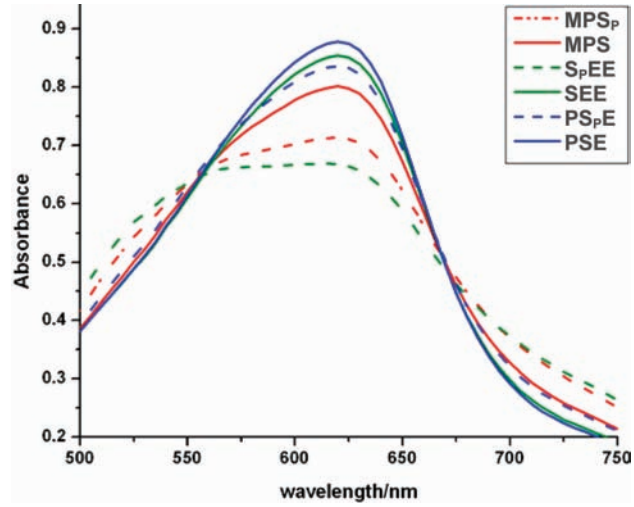


Figure 7. Example of spectral differences for one combination of host, metal, and indicator for the six target tripeptides. [**LG_1**] = [**IN4**] = [Cu^{II}] = 3×10^{-5} M, [target tripeptide] = 3×10^{-4} M, in water, buffered with Tris (50 mM) to pH 7.4, at 25 °C.

Before examining the differentiation of the peptides, it is instructive to take a look at some of the raw data that were ultimately processed by linear discriminant analysis. First, as an example of the differences found between the peptidic substrates, Figure 7 shows the absorbance spectra obtained for the six targets with receptor **LG_1**, in the presence of Cu^{II} and pyrocatechol violet (**IN4**). Substantial differences were found for each peptide; furthermore, the presence of isosbestic points is highly indicative of simple indicator displacement. Only the absorbance at 625 nm for each peptide was used in the ultimate pattern recognition (see below).

As another example, Figure 8 shows differences in the absorbance data for repetitions of several host/metal/indicator combinations with two different target tripeptides. The blue bars represent the differential signals obtained, i.e. the difference in absorbance obtained in the presence vs in the absence of the target tripeptides; larger blue bars characterize combinations that gave greater difference between the two tripeptides. The purple bars in the plot represent 3 times the standard deviations of the nine replicate measurements. Several combinations of peptide/metal/indicator can be found that give differences in differential signal for the two peptides that are greater than the error in the signals themselves (i.e., the difference between the blue bar value is greater than the sum of the purple bar values). These combinations are effectively responsible for differentiating the two tripeptides outside of the noise in the experimental data. As long as different combinations also gave differentiation among the other peptides, then a pattern recognition protocol would differentiate all the peptides; indeed, this turned out to be the case.

2.8. Pattern Recognition. The fingerprinting of the target tripeptides, both in their phosphorylated and nonphosphorylated forms as described above, produced a staggering amount of data, the sheer complexity of which makes it virtually impossible to

(29) Hoekstra, W. J. *Curr. Med. Chem.* **2001**, *8*, 715.

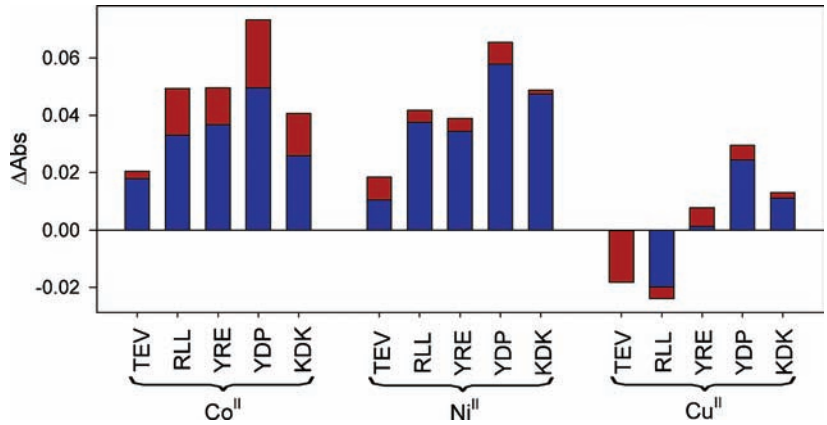


Figure 8. Example of differential response: absorbance changes for a series of [M^{II}(**LG**$_n$)] receptors (indicated by the one-letter-code amino acid sequence of their peptide arms) and celestine blue (**IN5**) in the interaction with a phosphorylated (pSEE) and nonphosphorylated tripeptide (SEE). [M^{II}] = [**IN5**] = [**LG**$_n$] = 3×10^{-5} M, [target tripeptide] = 3×10^{-4} M, in water, buffered with Tris (50 mM) to pH 7.4, at 25 °C. Presented results are an average of nine trials: the mean values (blue) and three standard deviations (purple) are shown.

find meaningful correlations by direct examination of the raw results. In order to reduce the dimensionality of the raw data set obtained from the array sensing, and to identify trends and patterns among the peptides, or between peptides and descriptor variables, computer-assisted pattern recognition algorithms were employed. Specifically, we constructed linear discriminant analysis models of the raw data set,[30] and of subsets thereof. Through statistical analyses, we intended to identify which components of the diversity that we introduced into our array play the most important role in discriminating the targets: is it the peptide arm sequence on receptors **2**, the choice of metal ions, or a particular indicator?

Linear discriminant analysis models were built utilizing a classical mode of variable selection. In short, discriminant functions Z_k were created (one less than the number of the predefined data classes), which are linear combinations of the descriptor variables with the most discriminating ability. They are of the form:

$$Z_k = a_1x_1 + a_2x_2 + a_3x_3 + ... + a_nx_n + C = \sum_{i=1}^{n} a_ix_i + C$$

where x_i = discriminating variables (i.e., the measured values of the independent variables used to predict group membership, e.g. ΔAbs in our case), a_i = discriminant weights (a measure of the discriminating/classifying ability of variable x_i), C = constant. A training set, consisting of experimental measurements for preclassified cases, enables the determination of the a_i weights for each data class. Once the a_i weights have been established, these discriminant scores (Z_k) can be used on new data, not included in the original training set, in order to predict class membership; in other words, the model is capable of classifying unknown analytes. Success in classification of analytes in discriminant analysis models is manifested graphically by large spatial separation of scores between analyte classes, and similar scores for experimental replicates within each analyte class. By using statistical classification methods, we were able to obtain specificity out of our series of unspecific, cross-reactive hosts.

The spectra obtained in this final study were analyzed in a fashion similar to that used for metal/indicator screening. For each metal/receptor/indicator combination, spectra obtained in the presence of the phosphorylated and nonphosphorylated tripeptides were compared, and the absorbance values of the bands showing the greatest difference between peptides were used in the final analysis. More specifically, the absorbance variables were introduced in a forward stepwise discrimination model performed using the XLSTAT program,[31] and the ones deemed to have the most discriminating power were selected on the basis of their F values (a measure of the total variance attributable to a given variable, or in other words, a measure of the overall impact of that specific variable in the model).[32]

In summary, in order to discriminate the three phosphorylated and three nonphosphorylated target peptides, constituting a total of six classification groups from the point of view of the discriminant analysis, we constructed five discriminant functions. These were obtained as linear combinations of data from 17 receptor/metal combinations, selected on the basis of their F values. Most of the total variance (99%) was captured by the first four functions.

Plots of the discriminant scores, Z_1 vs Z_2, and Z_1 vs Z_3, obtained for the replicate measurements, are shown in Figure 9. Excellent classification of all peptides is evident throughout, i.e. large spatial separation between tripeptides and clustering of the nine experimental replicates within each peptidic group. Z_1 discriminates all the peptides to varying extents: indeed, MPS and MPS_p are only discriminated along this axis. In contrast, Z_2 discriminates PSE and PS_pE, as well as SEE and S_pEE, but only to a smaller extent. Z_3 further differentiates MPS and MPS_p and SEE and S_pEE.

Validation of the forward discriminant model was carried out using three methods: the resubstitution method, the leave-one-out test, and random selection of five replicates from each peptidic group to create a discriminant model which was subsequently used to determine the identity of the four remaining experimental replicates from each peptidic group.[30] For the first two methods, 100% reclassification was obtained, whereas for the last validation method, only one experimental replicate was misclassified. These high rates of classification attest to the soundness of the training set.

(30) Otto, M. *Chemometrics: Statistics and Computer Application in Analytical Chemistry*, 2nd ed.; Wiley-VCH: Weinheim, 1998.

(31) As implemented in Addinsoft XLSTAT 2009 software package, http://www.xlstat.com.

(32) Hill, T.; Lewiki, P. *Statistics: Methods and Applications: A Comprehensive Reference for Science, Industry, and Data Mining*; StatSoft, Inc.: Tulsa, OK, 2006.

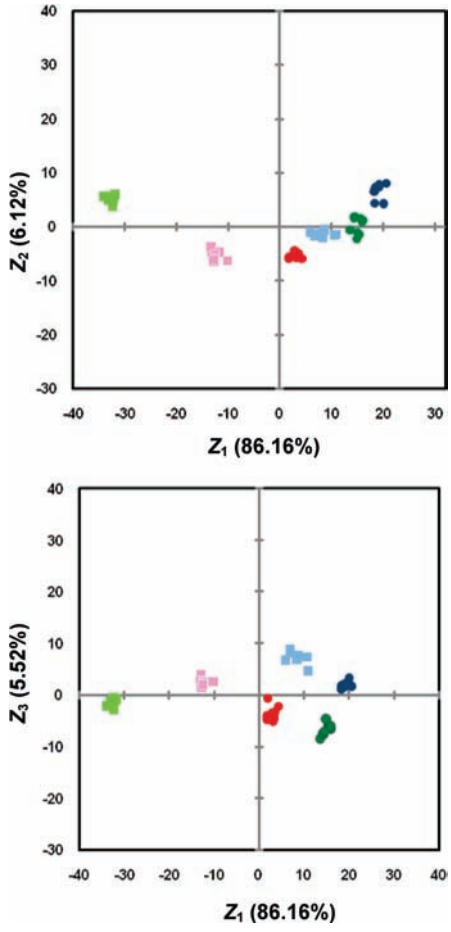


Figure 9. Plots of discriminant scores for all obtained data points, showing good spatial separation of different peptide classes and excellent clustering of experimental replicates. (red ●) MPS, (pink ■) MPS_P; (dark blue ●) PSE, (light blue ■) PS_PE; (dark green ●) SEE, (light green ■) S_PEE.

To gain insight into which combinations of peptidic arm sequences, metal ions and indicators in the array were most influential in discriminating the target tripeptides, classification models of reduced libraries were constructed. In our first reduced library, receptors composed of the same amino acid sequence, but different metal ions and indicators were used. When a single peptidic arm sequence was used, namely $\mathbf{LG_1}$, classification models indeed still featured discrimination of the peptides, but reduced spatial separation between peptides and reduced clustering of experimental replicates were found. The classification rates dropped to 88% (for the resubstitution method) and 86% (for the leave-one-out method). Similarly impaired discriminant score plots and classification rates for models constructed with the other four receptors ($\mathbf{LG_2}$−$\mathbf{LG_5}$) indicated that each receptor could indeed stand on its own in classifying the target tripeptides with nearly 90% confidence, but the variability in the peptidic arm sequences is very valuable in fully discriminating the target tripeptides.

Furthermore, a library constructed only with gallocyanine (**IN9**) outperformed similar libraries containing either indicator pyrocatechol violet (**IN4**) or celestine blue (**IN5**) in successfully classifying the target tripeptides. Classification rates were determined to be 100% for the resubstitution method and 92.6% for the leave-one-out method (four misclassifications).

Interestingly, a library containing all the peptides and indicators, but including only Cu^{II}, successfully discriminated the target tripeptides with 100% classification. On the other hand, such discrimination was not possible with arrays using only Ni^{II} or Co^{II}. However, even in the case of Cu^{II}, the classification plots clearly showed reduced analyte separation and worse clustering of experimental replicates.

The results summarized above make it clear that different peptidic arm sequences in the receptors and different indicators were critical for full discrimination of the peptides, whereas different metals were helpful but not nearly as critical. Differences in affinities between the indicators for the hosts must therefore impart differential binding affinities for the five $\mathbf{LG}_n$ receptors to the six tripeptide targets. Analogously, the peptide arms in [M^{II}(**2**)] must also make cross-reactive binding contact with the six tripeptide targets. Including different metal ions did increase the classification accuracy, but it turned out to be much less influential than originally anticipated.

3. Summary

This contribution reports on the ability to take a proven molecular recognition scaffold and diversify it to create patterns for a class of analytes. Specifically, a C_{3v} symmetric Cu^{II}-based phosphate receptor was elaborated into a library of receptors to target six different tripeptides that vary by sequence and phosphorylation state. Molecular diversity was created in the receptor library by varying a short sequence of amino acids appended to the receptor's arms and by varying the bound metal ion. For optical signaling, a series of indicators were used in indicator displacement assays, thereby introducing further diversity.

Extensive synthesis was employed to create the receptors, and prove their utility. Screening of the receptor library was performed prior to pattern analysis due to our past precedent that shows such screening improves the chemometric analysis.[19]

Linear discriminant analysis was applied to the optical signature arising from a set of five receptors, three metal ions, and three indicators upon exposure to the six peptide targets. Also, by restricting the LDA training sets and repeating the analysis, we were able to identify the key elements for creating cross reactivity in the array. It was found that the differential responses generated by varying the appended peptidic arms and indicators were critical to successful LDA classification. The study shows that, by making libraries of receptors that are biased toward classes of analytes, one can generate differential arrays for fingerprinting purposes.

Acknowledgment. We gratefully acknowledge financial support for the work described herein from the NSF (CHE-0716049) and the Welch Foundation.

Supporting Information Available: Detailed synthetic procedures, product characterization, spectral data, discriminant score plots, complete ref 18. This material is available free of charge via the Internet at http://pubs.acs.org.

JA9041675