Electronic Supplementary Material (ESI) for ChemComm.

Supporting information for

A cross reactive sensor array to probe divalent metal ions

A. M. Mallet,[a] A. B. Davis,[b] D. L. Davis,[b] J. Panella,[b] K. J. Wallace,[b]* and M. Bonizzoni[a]*

(a) Department of Chemistry, University of Alabama, Tuscaloosa, Alabama 35487, United States

(b) Department of Chemistry and Biochemistry, University of Southern Mississippi, Hattiesburg, Mississippi 39406 United States

General techniques: One-dimensional ^{1}H, and ^{13}C NMR spectra were recorded on a Bruker *Avance 400* spectrometer operating at a proton frequency of 400.13 MHz and equipped with a standard BFO 5 mm two channel probe in the appropriate deuterated solvents or were recorded on a Bruker *Ultrashield plus* 400 MHz spectrometer. Chemical shifts are reported in parts per million (ppm) downfield from tetramethylsilane (0 ppm) as the internal standard and coupling constants (*J*) are recorded in hertz (Hz). The multiplicities in the ^{1}H NMR spectra are reported as (br) broad, (s) singlet, (d) doublet, (dd) doublet of doublets, (ddd) doublet of doublet of doublets, (*t*) triplet, (*sp*) septet and (*m*) multiplet. All spectra were recorded at ambient temperature, unless otherwise stated. UV-Vis experiments were performed on a Beckman DU-800 UV-Vis spectrometer. Low resolution mass spectra were measured with Finnigan TSQ70. IR spectra were recorded on a Nicolet *Nexus 470* FT-IR paired with a *Smart Orbit* ATR attachment. The characteristic functional groups are reported in wavenumbers (cm^{-1}), and are described as weak (w), medium (m), strong (s) and very strong (vs). Fluorescence experiments were carried out on a QuantaMaster™ 40 Intensity Based spectrofluorometer from PTI technologies in the steady-state. High-resolution mass spectra (HRMS) were recorded at the Department of Chemistry of The University of Alabama using an *AutoSpec-Ultima* NT sector instrument.

Multivariate data techniques: spectroscopic data was acquired on a BioTek *Synergy II* multimode microwell plate reader, capable of measuring absorbance spectra (through a monochromator), and steady-state fluorescence intensity (through bandpass filter sets). The sample compartment in this instrument is electrically thermostatted. Experiments were laid out by hand using Eppendorf Research multichannel pipettors and disposable plastic tips into microwell plates with clear bottom for UV and fluorescence (Greiner BioOne), in 384-well configuration. The plates were made of non-treated (medium binding) polystyrene with black walls (to minimize scattered light) and clear flat bottoms. Each well invariably contained 300 μL of solution.

Synthesis and characterization:

General procedure for molecular probes (**2**)**,** (**3**) and (**5**). 7-(Diethylamino)-4-hydroxycoumarin (1.0 mmol), the appropriate primary amine (1.0 mmol), and triethylortho-formate (1.5 mmol) were refluxed in propan-2-ol (5 mL) for two hours. The reaction was allowed to cool to room temperature. The resulting solid was collected by vacuum filtration and washed with propan-2-ol. The characterization of probes **1** and **4** have been reported previously.[1]

Characterization of probe (2): Yield 224 mg, 0.66 mmol, 66% yield; ^{1}HNMR (400 MHz, $CDCl_3$) δ 13.57 (d, *J* = 12.2 Hz), 11.71 (d, *J* = 14.0 Hz), 9.60 (d, *J* = 13.6 Hz), 9.51 (d, *J* = 12.5 Hz), 8.48 – 8.39 (m), 7.95 (d, *J* = 9.0 Hz), 7.85 (d, *J* = 9.0 Hz), 7.74 (ddd, *J* = 9.4, 6.9, 1.8 Hz), 7.20 – 7.11 (m), 7.07 (dd, *J* = 24.0, 8.1 Hz), 6.56 (dd, *J* = 9.0, 2.4 Hz), 6.36 (d, *J* = 2.4 Hz), 3.43 (q, *J* = 7.1 Hz), 1.23 (t, *J* = 7.1 Hz); ^{13}C NMR (100 MHz, $CDCl_3$) δ 182.2, 181.3, 177.6, 165.9, 164.2, 157.3, 156.9, 153.1, 153.0, 152.8, 151.2, 149.2, 138.8, 128.3, 127.5, 121.2, 113.0, 112.6, 108.7, 108.5, 108.4, 99.3, 97.2, 97.1, 44.9, 12.5; IR (ATR solid); 3060 (w) ν_{NH}, 2974 (w) ν_{CH}, 1714 (s) ν_{CO} (delta lactone), 1562 ν_{CO} (ketone) cm^{-1}; HRMS: $[M]^{+\bullet}$: Calc for $C_{19}H_{19}N_3O_3$ = 337.1426; found for $C_{19}H_{19}N_3O_3$ = 337.1431 and $[M\text{-}CH_3]^{+\bullet}$ Calc for $C_{18}H_{16}N_3O_3$ = 323.1192; found for $C_{17}H_{15}N_4O_3$ = 323.1197

Characterization of probe (3): Yield 155.6 mg, 0.46 mmol, 46% yield; ^{1}H NMR (400 MHz, $CDCl_3$) δ 13.71 (d, *J* = 12.6 Hz, 1H), 11.64 (d, *J* = 13.0 Hz, 1H), 8.92 (d, *J* = 14.0 Hz, 1H), 8.77 (d, *J* = 12.9 Hz, 1H), 8.64 (dd, *J* = 8.0, 2.5 Hz, 1H), 8.51 (d, *J* = 4.1 Hz, 1H), 7.92 (d, *J* = 9.0 Hz, 1H), 7.84 (d, *J* = 9.0 Hz, 1H), 7.67 (t, *J* = 11.3 Hz, 1H), 7.39 (dd, *J* = 8.2, 4.7 Hz, 1H), 6.57 (dd, *J* = 9.0, 2.2 Hz, 1H), 6.35 (d, *J* = 2.1 Hz, 1H), 3.43 (q, *J* = 7.0 Hz, 5H), 1.23 (t, *J* = 7.1 Hz, 7H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 181.1, 164.1, 157.3, 157.0, 153.5, 153.2, 152.0, 147.6, 147.5, 140.7, 140.6, 135.0, 128.2, 127.5, 124.6, 124.3, 124.3, 108.7, 108.6, 108.6, 108.5, 99.8, 99.3, 97.2, 77.3, 77.0, 76.7, 44.9, 12.5. IR (ATR solid); 3066 (w) $\nu_{NH,}$ 2966 (w) ν_{CH}, 1711 (s) ν_{CO} (delta lactone), 1608 ν_{CO} (ketone) cm^{-1}; HRMS: $[M]^{+\bullet}$: Calc for $C_{19}H_{19}N_3O_3$ = 337.1426; found for $C_{19}H_{19}N_3O_3$ = 337.1430 and $[M\text{-}CH_3]^{+\bullet}$ Calc for $C_{18}H_{16}N_3O_3$ = 323.1192; found for $C_{17}H_{15}N_4O_3$ = 323.1191.

Characterization of probe (5). Yield 346.4 mg, 1.02 mmol, 51% yield; ^{1}H NMR (400 MHz, $CDCl_3$-*d*): δ13.20 (d, 1H, *J* = 12.2 Hz, NH), 11.60 (d, 1H, *J* = 12.8 Hz, NH), 9.62 (d, 1H, *J* = 13.6 Hz, $CH_{enamine}$), 9.47 (d, 1H, *J* = 12.7 Hz, $CH_{enamine}$), 8.64 (t, 2H, *J* = 5.7 Hz, $CH_{aromatic}$), 7.92 (dd, 1H, *J* = 24.5, 9.0 Hz, $CH_{coumarin}$), 7.12 (t, 1H, *J* = 4.8 Hz, $CH_{aromatic}$), 6.57 (dd, 1H, *J* = 9.0, 2.1 Hz, $CH_{coumarin}$), 6.35 (d, 1H, *J* = 2.1 Hz, $CH_{coumarin}$), 3.44 (q, 4H, *J* = 7.0 Hz, CH_2), 1.24 (t, 6H, *J* = 7.1 Hz, CH_3). ^{13}C NMR (300 K, $CHCl_3$-*d*, 100 MHz): 202.8, 181.0, 164.0, 158.7, 158.7, 157.3, 156.7, 153.3, 152.8, 128.5, 128.0, 117.9, 117.9, 108.8, 108.5, 100.8, 97.2, 97.1, 44.9, 12.5. IR (ATR solid); 3062 (w) $\nu_{NH,}$ 2968 (w) ν_{CH}, 1713 (s) ν_{CO} (delta lactone), 1601 and 1551 ν_{CO} (ketone) cm^{-1}; HRMS: $[M]^{+\bullet}$: Calc for $C_{18}H_{18}N_4O_3$ = 338.1379; found for $C_{18}H_{18}N_4O_3$ = 337.1375 and $[M\text{-}CH_3]^{+\bullet}$ Calc for $C_{17}H_{15}N_4O_3$ = 323.1140; found for $C_{17}H_{15}N_4O_3$ = 323.1154.

UV-Vis studies: A stock solution of molecular probes **1** and **5** (0.31 mM) was prepared by dissolving the appropriate weight in 20 mL DMSO. From this stock solution a 0.31 mM solution of probe **5** was prepared in a one mL quartz cell for UV-vis spectroscopic studies. A zinc chloride stock solution (0.31 mM) was prepared in DMSO. Aliquots of the metal salt solution (5 μL) were added to the cell (each addition ≡ 0.1 equivalents of Zn(II) cation).

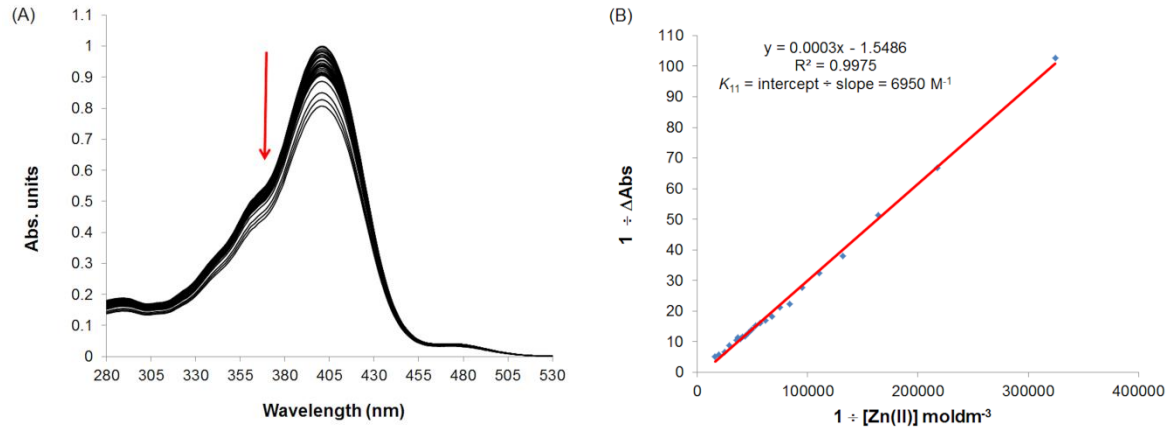


***Figure S1*:** (A) UV-vis titration of molecular probe **1** upon the addition of Zn(II)chloride and (B) the Benesi-Hildebrand plot: K_{11} = 6950 at 298 K.

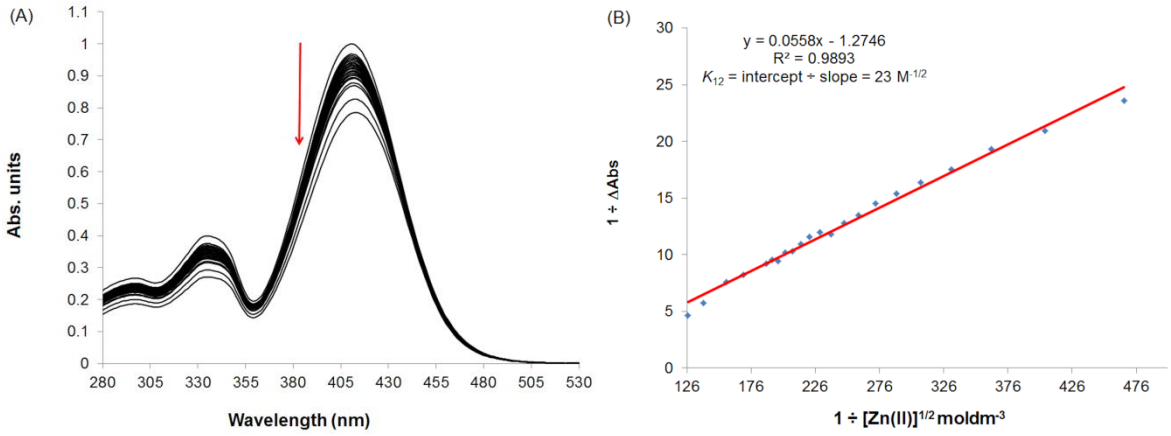


***Figure S2*:** (A) UV-vis titration of molecular probe **5** upon the addition of Zn(II)chloride and (B) the Benesi-Hildebrand plot: K_{12} = 23 at 298 K.

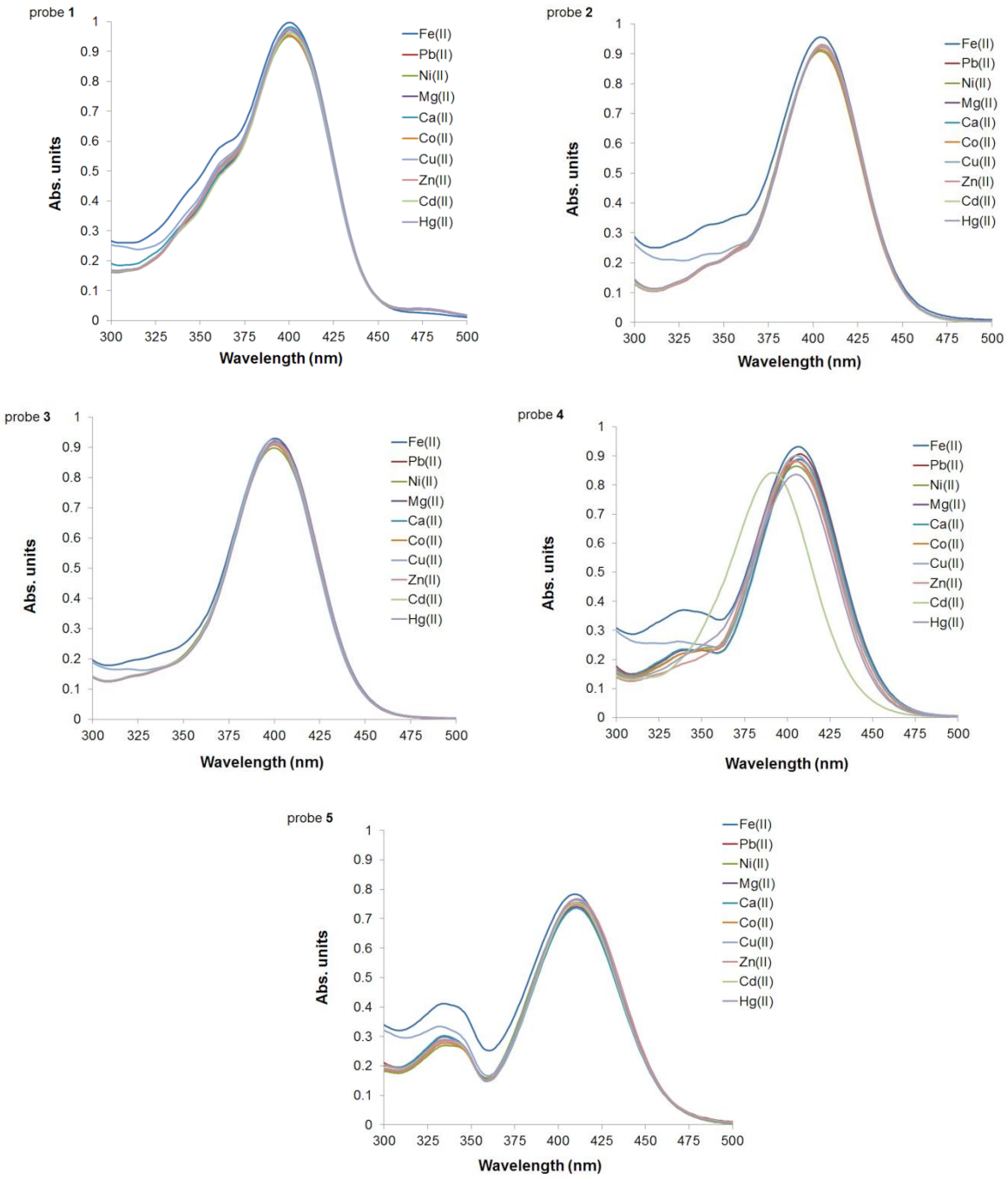


***Figure S3*:** Absorbance spectra of probes **1** to **5** (16 μM), upon the addition of different MCl_2 (32 μM)

Fluorescence studies: For all fluorescence studies stock solutions were prepared as described in the UV-vis section. Compounds **1** and **5** were then diluted by removing 100 μL and diluting to 2 mL to produce a final concentration of 16 μM. Aliquots of the metal salt solution (10 μL) were added to the solution of ligand (each addition ≡ 0.1 equivalents of Zn(II) cation). Excitation wavelength 408 nm, slit widths 0.35 mm scanned from 420 to 780 nm.

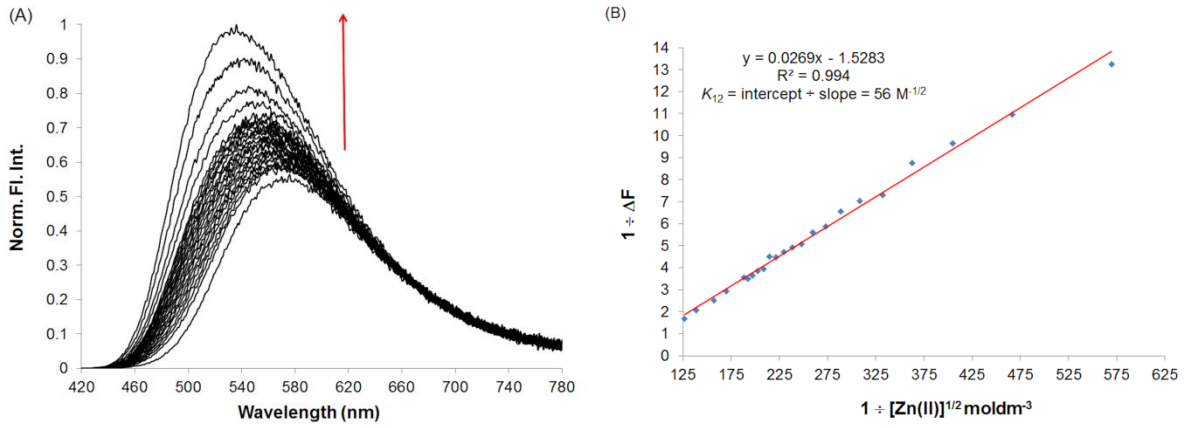
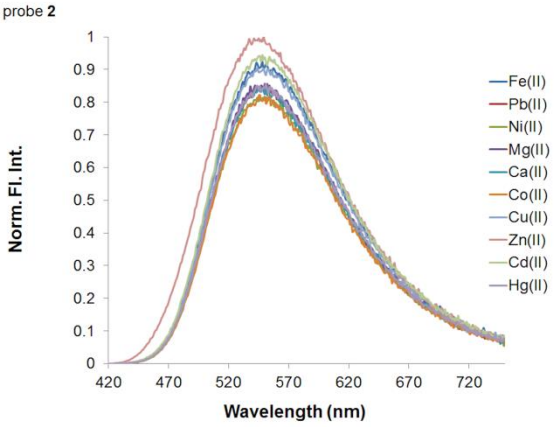


***Figure S4*:** (A) Fluorescence titration of molecular probe **1** upon the addition of Zn(II)chloride and (B) the Benesi-Hildebrand plot: K_{12} = 6950 at 298 K.



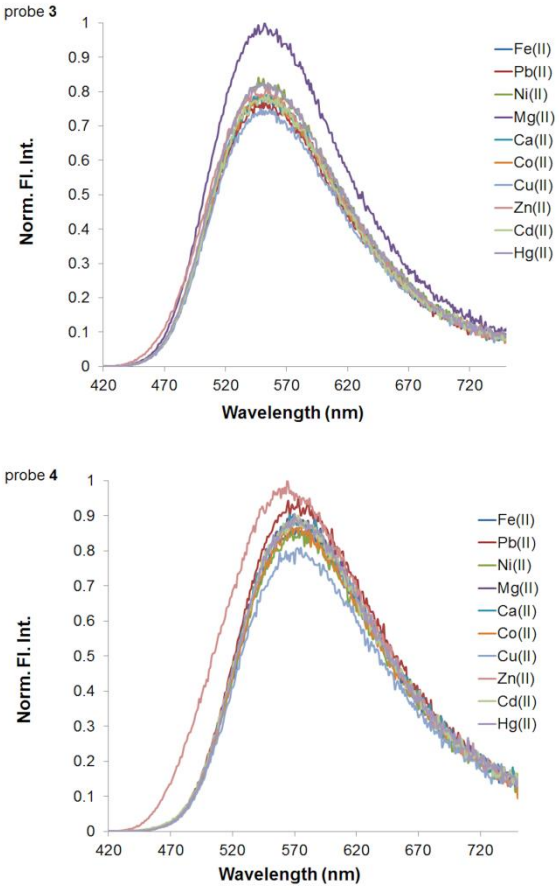


***Figure S5*:** Normalized fluorescence spectra of probes **2, 3** and **4** (16 μM, λ_{ex} = 408 nm), upon the addition of different MCl_2 (32 μM, λ_{ex} = 408 nm).

Determination of the univariate limit of detection

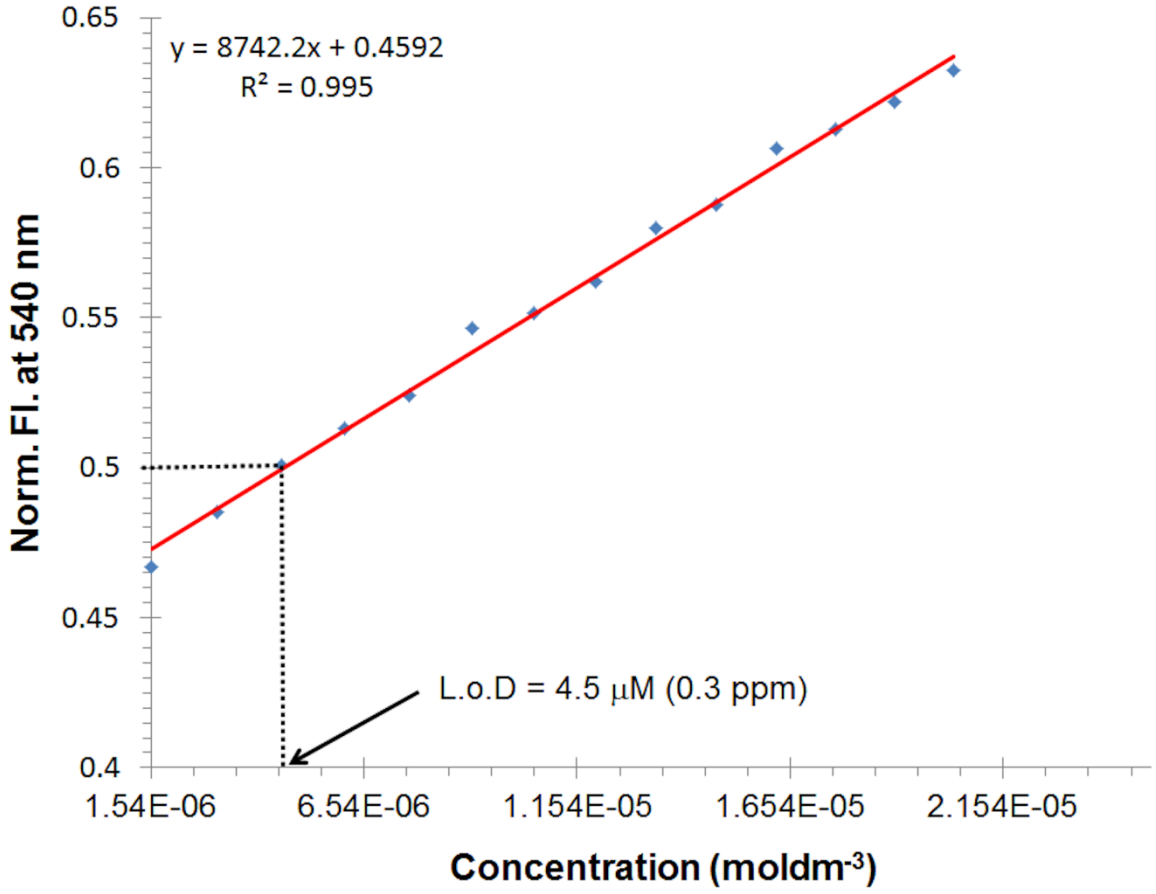


***Figure S6*:** To determine the limit of detection (LoD), between probe **5** and Zn(II) the method of least squares was used to give a line of regression. The confidence limit of the slope is defined as b ± t sb, where t is the t-value taken from the desired confidence and n-2 degrees of freedom. A 95% confidence level (t-value 2.14, df = 14).

Multivariate experimental set-up: All experiments were carried out in dimethylsulfoxide (DMSO). Experimental temperature was thermostatted internally to 24°C. Plates were read in a multimode plate reader immediately after preparation. For the preparation of the training set described here, manual dispensing of solutions into each plate generally required 3-4 hours; reading time typically required 30-45 minutes per plate. In that time, we did not observe any significant evaporation, so we could afford not to seal the plates, which might have otherwise impacted the sensitivity of the measurement. It is important to note that the training set need only be measured every once in a while, similarly to a calibration curve in a univariate analysis. The time required to measure a sample once the training set has been acquired is much less than described here (on the order of one minute per each different sample, i.e. a few seconds per replicate within each sample).

For a typical multivariate binding experiment, a series of two 384-well plates were utilized. Each plate was laid out to contain the following: two probes among the coumarin-enamine compounds, 10 metal chloride analytes with each metal ion having 18 replicates; 12 replicates of DMSO (used for blanking), 12 replicates of probe **1** (used for normalization of data across plates). A schematic of the plate layout using probes **2** and **3** is illustrated in figure S7. The probe concentration was kept constant at 1.6×10^{-5} M (16 µM). The metal chloride concentration was at 4.8×10^{-5} M (48 µM). A 3-to-1 ratio of metal-to-probe was chosen as no further spectral change was seen after the three equivalents of metal.

Rows	**Columns 1-6**	**Columns 7-12**	**Columns 13-18**	**Columns 19-24**
A - C (probe **2**)	Mg	Ni	Zn	Fe
D – F (probe **2**)	Ca	Cu	Pb	Co
G – I (probes **2** and **3**)	Cd (probe **2**)	Hg (probe **2**)	Mg (probe **3**)	Ni (probe **3**)
J – L (probe **3**)	Zn	Fe	Ca	Cu
M – O (probe **3**)	Pb	Co	Cd	Hg
P	Probe **1** control = 12 replicates		DMSO blank = 12 replicates	

Figure S7: Typical layout of a multivariate array sensing plate.

Multivariate analysis was based on 10 instrumental variables per sensor: fluorescence intensities were collected in the following channels ($\lambda_{ex}/\lambda_{em}$): 330/450 nm, 330/528 nm, 330/580 nm, 380/450 nm, 380/528nm and 380/580 nm; absorbance values were collected at the following wavelengths: 330, 380, 400, and 430 nm. The overall raw data matrix consists of 180 rows (analyte replicates) by 40 columns (variables). A total of 7200 raw data points was collected as depicted in Figure S8.

10 metal analytes	18 data points per analyte	= 180 total data points
4 sensors	10 variables per sensor	= 40 total variables
7200 raw data points		

Figure S8: Schematic of raw data point acquisition.

Multivariate data analysis: In order to evaluate the discriminatory power of the molecular probes, Linear Discriminant Analysis (LDA) was used as a multivariate approach to evaluate the datasets. LDA is a statistical treatment that is used for reinterpretation of a multidimensional data set. All multivariate analyses were performed in the commercial *Mathematica* program (release 10.1) published by Wolfram Research Inc..

As a first step towards data analysis, the raw experimental data was normalized to the corresponding values measured for sensor **1**, to ensure experimental consistency among the two microwell plates forming an experimental data set. Probe **1** does not possess a chelating binding site on the aromatic ring, but it contains the same chromophoric / fluorogenic unit as the other compounds, therefore it was used as a control to provide plate-to-plate consistency to the system. Normalization was achieved by dividing the response of each sensor in the panel by the average response measured for sensor **1** in the corresponding channel as illustrated in figure S9.

$$\text{7200 raw data points} \quad \rightarrow \quad \frac{raw\ data\ read}{averaged\ probe\ \mathbf{1}\ signal} \rightarrow \quad \text{7200 normalized data points}$$

Figure S9: schematic for normalization of multivariate data analysis.

After normalization of the data, potential outliers due to obvious gross errors during the plate layout were then removed. The 18 replicate points obtained from each metal chloride were subjected to principal component analysis (PCA) to estimate their dispersion. Points well outside a 95% confidence interval (CI) were considered outliers and removed. As an example, see figure S10 for the Ni(II) analyte: in this case, data point five lies well outside the 95% CI and therefore was removed.

PCA analysis of the single analyte replicates was found to be the quickest way to spot similar problematic points that would have otherwise significantly skewed the results of further analysis. Such points would be very difficult, if not impossible, to spot by eye on the small 384-well plates. Similar problematic points were removed from the nine remaining metal chloride sets. Consequently outliers were removed in a very conservative way, and only when, using the guidance of the PCA results, some damage or blemish was found upon careful inspection of the physical plate. A total of 30 data points were removed from the original 7200 points before further analysis.

After the removal of outliers, the resulting data set was subjected to LDA. The loading plot depicted in Figure S11 reports on the contribution of each one of the 40 instrumental variables to

the overall discriminatory power of the sensing system. The loadings report on the contribution of each original instrumental variable to the two factors we selected. From the loading plot, it was observed that the fluorescence intensities at 330/450 nm and 380/450 nm for sensor **4** are the most important contributors, as well as the absorbance values at 400 nm and 430 nm for sensor **5**. This in turn indicates that these two probes, along these channels, were particularly valuable for the discrimination.

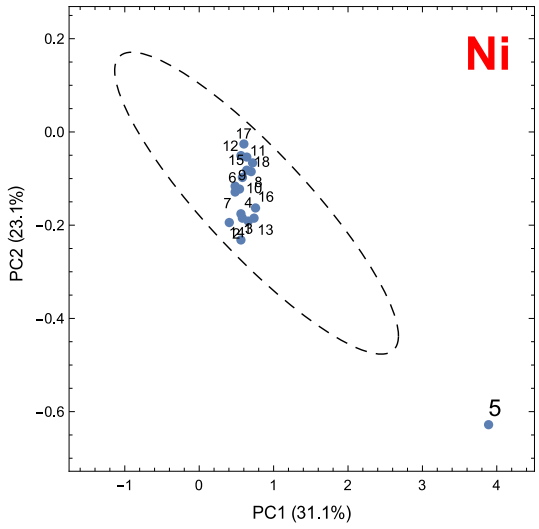


Figure S10: PCA score plot of the nickel chloride replicate set, with 95% confidence ellipsoid shown as a dashed line. Point 5 is clearly outside the confidence interval, and it was separately confirmed to be a faulty data point (see text), so this point was removed from the dataset before further analysis.

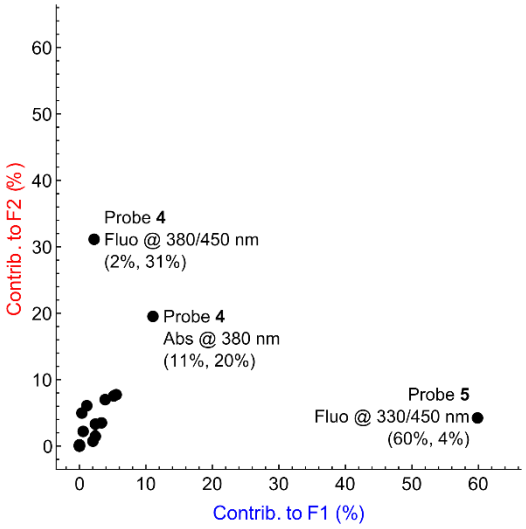


Figure S11: Loadings plot corresponding to the linear discriminant analysis (LDA) presented in the main manuscript. The most important contributor to discrimination along factor 1 is the fluorescence signal from sensor 5 upon excitation at 330 nm and detection at 450 nm; similarly, the most important contributor to the discrimination along the second factor is a fluorescence signal arising from sensor **4**. Minor contributions are also present from the other sensors and variables in the array: labels have been omitted for clarity.

Hierarchical clustering analysis on the full data set was finally carried out in order to capture as much information as possible. The analysis was carried out using Manhattan intercluster distance and Ward linkages and resulted in the dendrogram reported in Figure 5.[2] This method was used to test the classification capability of the system at hand and to produce the corresponding mis-classification matrix. Clustering analysis was carried out using Wolfram Research's *Mathematica*, using a built-in clustering and classification routines.

The mis-classification matrix was constructed from a naïve Bayes classifier.[3] The full data set was divided in two parts: two thirds of the data were used as a training set, whereas the remaining third was used as a test set for the classifier. Six out of 18 replicates were drawn randomly from each metal ion set to provide test samples; the remaining points provided the training set.

References:

1. A. B. Davis, R. E. Lambert, F. Fronczek, P. J. Cragg and K. J. Wallace, "An activated coumarin-enamine Michael acceptor for CN-" New J. Chem, 2014, 38, 4678-4683.

2. J. C. Miller and J. N Miller *Statistics for Analytical Chemistry* (3rd ed.). Ellis Horwood, **1993**.

2. S. Russell; P. Norvig. *Artificial Intelligence: A Modern Approach* (2nd ed.). Prentice Hall **2003**. ISBN 978-0137903955.