SUPPORTING INFORMATION

DNA Polyfluorophores as Highly Diverse Chemosensors of Toxic Gases

Chi-Kin Koo, Florent Samain, Nan Dai and Eric T. Kool*
Department of Chemistry, Stanford University, Stanford, California 94305-5080
*to whom correspondence should be addressed: kool@stanford.edu

CONTENTS:

Experimental

General Materials and Methods. Solvents and chemicals from commercial sources were used as received. ^{1}HNMR, ^{13}C-NMR, ^{31}P-NMR and ROESY NMR spectra were taken on either a 400 MHz Varian Mercury, 500 MHz Varian Inova, or 600 MHz Varian Inova spectrometer. Chemical shifts are reported in ppm with the solvent resonance as the internal standard. ^{31}P NMR spectra were referenced to 85% H_3PO_4 in water as external standard (0.0 ppm). ESI-MS was taken on a Finnigan LCQ Mass-spectrometer. HR-MS was performed on a Micromass Q-Tof hybrid quadrupole-time of flight LC-MS at the SUMS facilities at Stanford University. MALDI-TOF MS were performed on a Voyager-DE RP Biospectrometer at the PAN facilities at Stanford University. Absorption spectra were recorded on Cary 1 UV-Vis spectrometer.

Monomer and Library Syntheses. The **Y**, **E** and **D** monomer deoxyribosides (as α anomers) were prepared as described previously.[1] The α-porphyrin nucleoside (**1**), its 4,4'-dimethoxytrityl (**2**) and α-porphyrin phosphoramidite (**3**) derivatives were synthesized according to the published procedures[2] from 2,5-anhydro-3-deoxy-4,6-di-O-toluoyl-D-ribohexose[2] and dipyrromethane[3]. In this study only the α anomer of the free nucleoside was collected, with structure confirmed by NOESY NMR experiments (see below). Altogether, four fluorescent nucleosides (**Y**, **E**, **D** and **P**) were included to yield 256 tetramer-length ODF sequences in all combinations. The ODF library was constructed on 130 μm amine-functionalized polyethylene glycol-polystyrene (PEG-PS) beads (NovaSyn TG amino resin (Novabiochem; average loading: 100 mmol/g) using standard tagging procedures[4] and split-and-mix methods[5] as previously reported. Zn(II) complexation of the porphyrin-containing sequences was performed after the bead-supported library was built, by incubating with excess zinc(II) trifluoromethanesulfonate ($ZnOTf_2$) in ethanol at 25 °C for 12 hours. The successful incorporation of zinc(II) was later conformed by high resolution mass spectrometry (HRMS) and UV-Vis absorption spectroscopy (see below). Deprotection of the phosphate backbone was then performed in concentrated ammonia solution (12 hours at 55 °C) to yield the library used for screening.

Analyte Preparation. To control the concentration of gas analytes used for the cross-testing experiments, gases were first flushed into a 50 mL closed round-bottom flask equipped with a septum and equilibrated with an oil bubbler. To dilute these, 5 mL of each gas was collected with a gas-tight syringe (Hamilton) and injected into a 50 mL round-bottom flask which had been previously flushed with dry argon gas, yielding a tenfold dilution. A volume of 35 μL of diluted gas was then collected with a gas-tight syringe and injected into a screw-cap quartz cuvette (Starna Cells), which has a volume of 3.5 mL of atmospheric air and was equipped with a silicone rubber septum, to make up a final concentration of 1000 ppm of gas (±10%) inside the cuvette.

Screening, Image Processing and Decoding. The initial screening process was

performed in a closed quartz fluorescence cell equipped with a septum, in which beads from the library were placed onto a small microscope slide. The beads were exposed to approximately 1% of each toxic gas in air as described above. Fluorescence was monitored under an epifluorescence microscope (Nikon Eclipse E800) with excitation at 340-380 nm, and all visible emission was observed (long-pass filter, >420 nm). Fluorescence images were taken with a Spot RT digital camera and Spot Advanced Imaging software before and after exposure to the gas for 2, 7, and 15 min in the chamber at room temperature. To determine emission changes in response to the gaseous analytes, we first constructed graphical 50% gray-based difference maps of the beads by inverting color/intensity of the image before exposure (i.e. making a photonegative) and merging it with the image taken after 15 min of exposure using 50% transparency (Adobe Photoshop, version 10.01). An example is shown in Fig. S6. Any part of the image that is 50% gray indicates no change, whereas beads that are darker than the 50% gray background reveal quenching, brighter beads show emission enhancement, and colors reflect a combination of the original ODF emission color and any wavelength shifts that occur during the sensing events. Beads that showed the strongest responses were picked up with a flame-pulled pipet and transferred into a capillary tube for sequence-decoding by electron-capture gas chromatography (EC-GC).[22]

Oligodeoxyfluoroside Synthesis. The 15 sensor ODFs selected from the initial screening were individually resynthesized on an Applied Biosystems 394 DNA/RNA synthesizer on a 1 μmole scale and possessing a 3'-phosphate group for characterization off of the beads. Standard β-cyanoethyl phosphoramidite chemistry was employed for coupling, but with an extended coupling time (999 s) for the non-natural nucleotides. Overall coupling efficiencies exceeded 70%. Similar to the library synthesis, Zn(II) complexation for porphyrin-containing sequences was performed after solid-phase ODF synthesis as described above. Oligomers were then deprotected and cleaved from the solid support with concentrated ammonia solution (55 °C, 12 h). The oligomers were purified by HPLC (Shimadzu 10 Series with a Jupiter 5u C5 column) using water (TEAA, pH~7.2) and acetonitrile as the solvent system. The purified oligomers were characterized by HRMS or MALDI-MS (see SI). Standard automatic DNA synthesis techniques were used except that 50/50 of cleavable controlled pore glass (CPG) and non-cleavable PEG-PS beads were included in the column, so that sensor beads and corresponding ODFs on standard cleavable CPG were made at the same time. This allowed the characterization of the ODFs off of the beads (Table S1, Fig. S7) while making the corresponding sensors on beads in the same synthesis.

Cross-screening and Chemometric Analysis. Cross-screening experiments were performed under an epifluorescence microscope with a 4X objective and fluorescence images were taken before and after the exposure of each analyte. The emission was then quantified as RGBL (red, green, blue, luma) values by using Adobe Photoshop. Average RGBL values

were determined within a 16×16-pixel box in the center of each bead picked from the images recorded before and after exposure. Color change, expressed in ΔR, ΔG, ΔB and ΔL, was calculated and averaged (5 beads per analyte for a given sensor). A set of quantitative color-change profiles was then constructed by plotting the ΔR, ΔG, ΔB and ΔL values for each analyte. Standard deviations were determined from the variance in the 5 beads to evaluate the accuracy and reproducibility of the responses. To analyze the response data quantitatively, principal component analysis (PCA) and agglomerative hierarchical clustering (AHC) analyses were performed with XLSTAT software (Addinsoft) using the ΔR, ΔG, ΔB and ΔL values as input. PCA allows the scattering of the data to be quantitatively analyzed by plotting the responses on new (non-Euclidean) axes that represent the most orthogonal components. To examine relationships between the types of responses of the 15 sensors, AHC (an agglomeration method based on the squared Euclidean distance and unweighted pair-group average linkage between centroids of the clusters) was carried out, generating dendrograms in which classes of sensor sequences can be identified according to their response patterns.

1',2'-Dideoxy-1'-(5-(15-phenyl) porphyrinyl)-α-D-ribofuranose. (1)

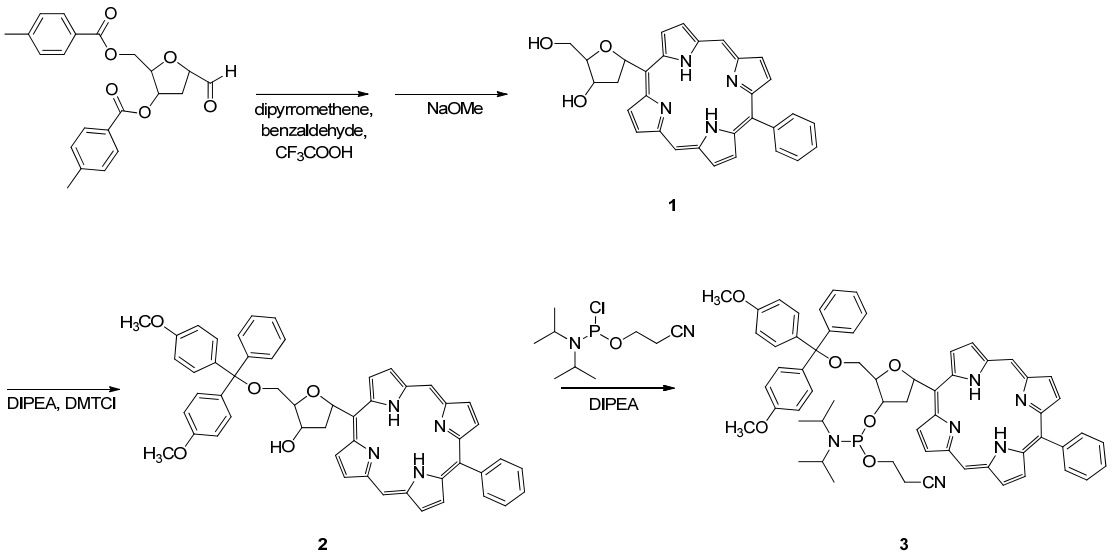
^{1}H-NMR (Pyridine-d_5, 500MHz): δ= 10.48 (2H, s), 10.30 (2H, d, J = 4.5Hz), 9.57 (2H, d, J = 4.5Hz), 9.51 (2H, d, J = 4.5Hz), 9.11 (2H, d, J = 4.5Hz), 8.27-8.29 (2H, m), 8.22 (1H, dd, J = 6.5Hz and J = 11Hz), 7.80-7.83 (3H, m), 5.71-5.76 (1H, m), 5.57-5.60 (1H, m), 4.64-4.71 (2H, m), 3.75-3.80 (1H, m), 3.62-3.80 (1H, m), -2.45 (1H, s), -2.52 (1H, s); ^{13}C-NMR (Pyridine-d_5, 125.7MHz) δ = 147.5, 147.4, 145.3, 145.1, 144.9, 141.4, 132.9, 132.4, 131.1, 130.7, 129.8, 129.6, 128.2, 127.6, 119.7, 118.2, 105.8, 88.8, 81.9, 74.2, 63.5, 52.5; ESI-MS calcd for $C_{31}H_{27}N_4O_3$ [$(M+H)^+$] 503.57, found 503.22

1',2'-Dideoxy-5'-(4,4'-dimethoxytrityl)-1'-(5-(15-phenyl) porphyrinyl)- α-D-ribofuranose. (2)

^{1}H-NMR ($CDCl_3$, 400MHz): δ= 10.30 (2H, s), 9.89 (2H, d J = 4.5 Hz), 9.45 (2H, d, J = 4.5Hz), 9.37 (2H, d, J = 4.5Hz), 9.07 (2H, d, J = 4.5Hz), 8.25-8.27 (2H, m, Ph), 7.93 (1H, dd, J = 6Hz and 10.4Hz), 7.79-7.82 (3H, m), 7.69-7.72 (2H, m), 7.56-7.60 (4H, m), 7.41 (2H, t, J = 7.2Hz), 7.31 (1H, t, J = 7.2Hz), 6.90-6.95 (4H, m), 5.13-5.11 (1H, m, H3'), 4.6-4.58 (1H, m, H4'), 5.16-5.22 (1H, m), 5.04-5.09 (1H, m), 3.80-3.84 (8H, m), 3.54-3.6.62 (1H, m), 3.21-3.30 (1H, m); ^{13}C-NMR ($CDCl_3$, 125.7MHz): δ = 158.6, 147.0, 146.8, 145.2, 144.7, 144.2, 141.0, 136.3, 136.2, 134.8, 132.5, 131.6, 130.8, 130.3, 128.6, 128.4, 128.1, 127.7, 127.1, 127.0, 119.4, 116.1, 113.3, 105.1, 87.0, 86.5, 82.5, 75.5, 66.0, 53.8, 51.71; ESI-MS calcd for $C_{52}H_{45}N_4O_5$ [$(M+H)^+$] 805.94, found 805.50.

1',2'-Dideoxy-5'-(4,4'-dimethoxytrityl)-1'-(5-(15-phenyl) porphyrinyl)-α-D-ribofuranose-3'-cyanoethyl-N,N-diisopropylphosphoramidite. (3)

^{1}H-NMR ($CDCl_3$, 400MHz): δ= 10.27 (2H, s), 9.90 (2H, d, J = 4.8 Hz), 9.41 (2H, d, J = 4.8 Hz), 9.35 (2H, d, J = 4.8 Hz), 9.04 (2H, d, J = 4.8 Hz), 8.25-8.23 (2H, m, Ph), 8.00-7.97 (1H, m), 7.80-7.61 (6H, m), 7.60-7.56 (3H, m), 7.37 (2H, q, J = 7.2 Hz), 7.28 (1H, t, J = 6.8 Hz), 6.91-6.83 (4H, m, DMT), 5.20-5.30 (1H, m), 5.13-5.18 (1H, m), 3.80-3.88 (2H, m), 3.78-3.75 (6H, m), 3.67-3.72 (4H, m), 3.39 (1H, m), 2.65 (1H, m), 2.5 (1H, t, J = 6.4 Hz), 2.43 (1H, q, J = 6.4 Hz), 1.33-1.20 (12H, m); ^{31}P-NMR ($CDCl_3$, 202MHz): 149.64, 150.23; HRMS calcd for $C_{61}H_{62}N_6O_6P$ [$(M+H^+)$] 1005.4468, found 1005.4470.



Scheme S1. Synthesis of α-porphyrin phosphoramidite (**3**).

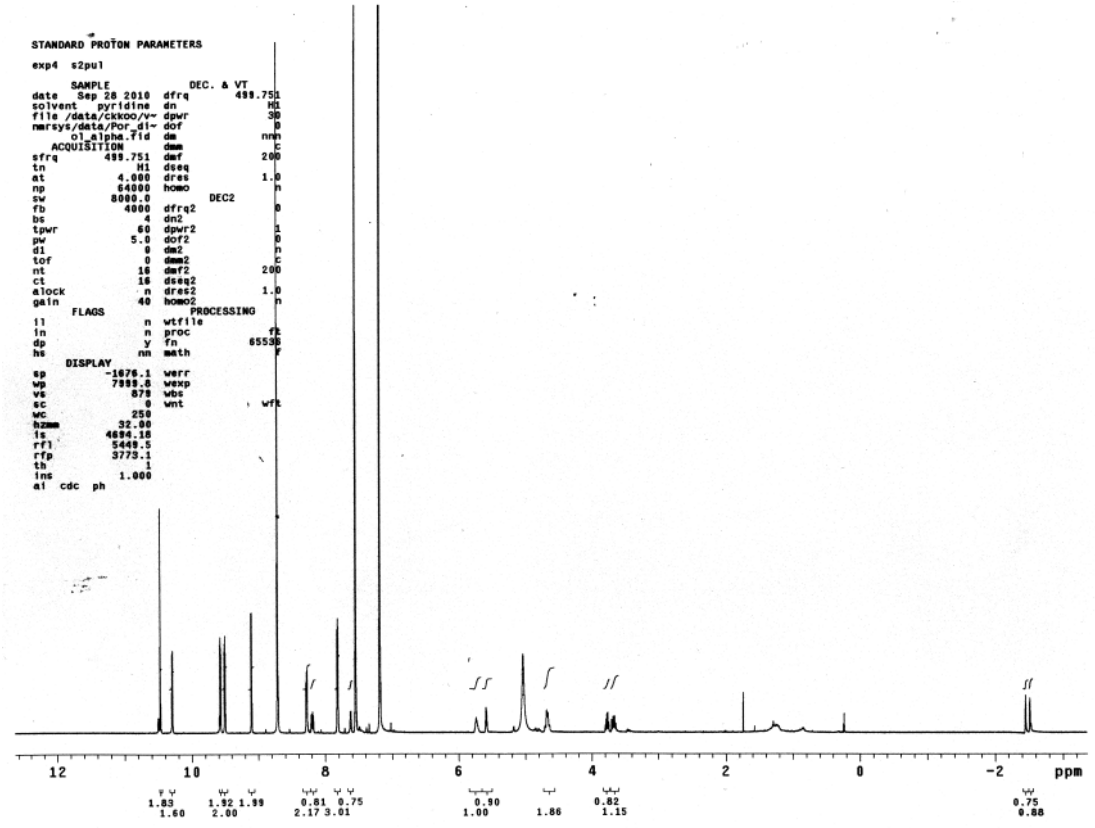


Figure S2. ^{1}H-NMR spectrum of **1** (pyridine-d_5).

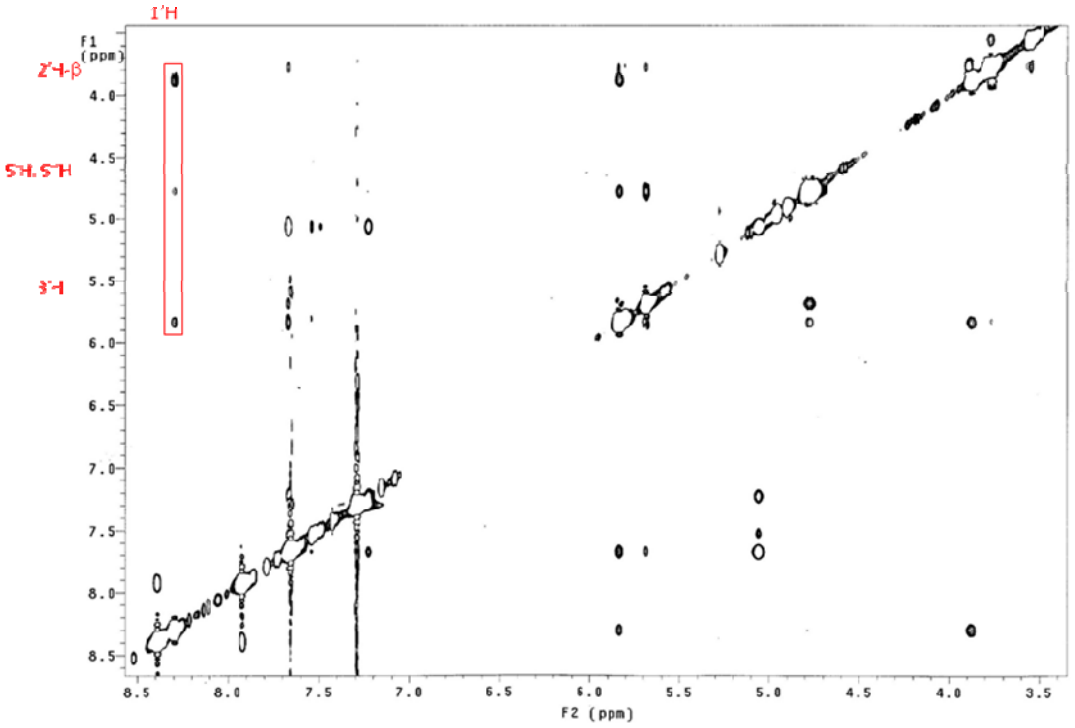


Figure S3. ROESY NMR spectrum of **1** (Pyridine-d_5).

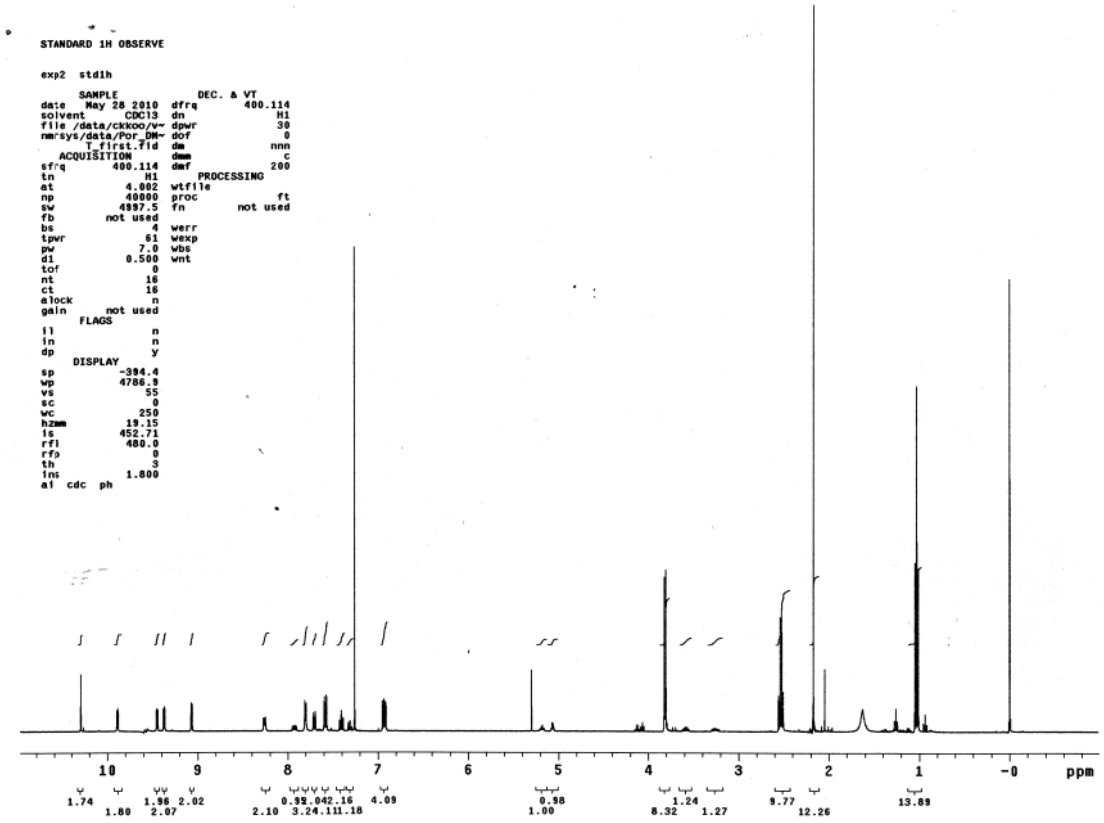


Figure S1. ^{1}H-NMR spectrum of **2** ($CDCl_3$).

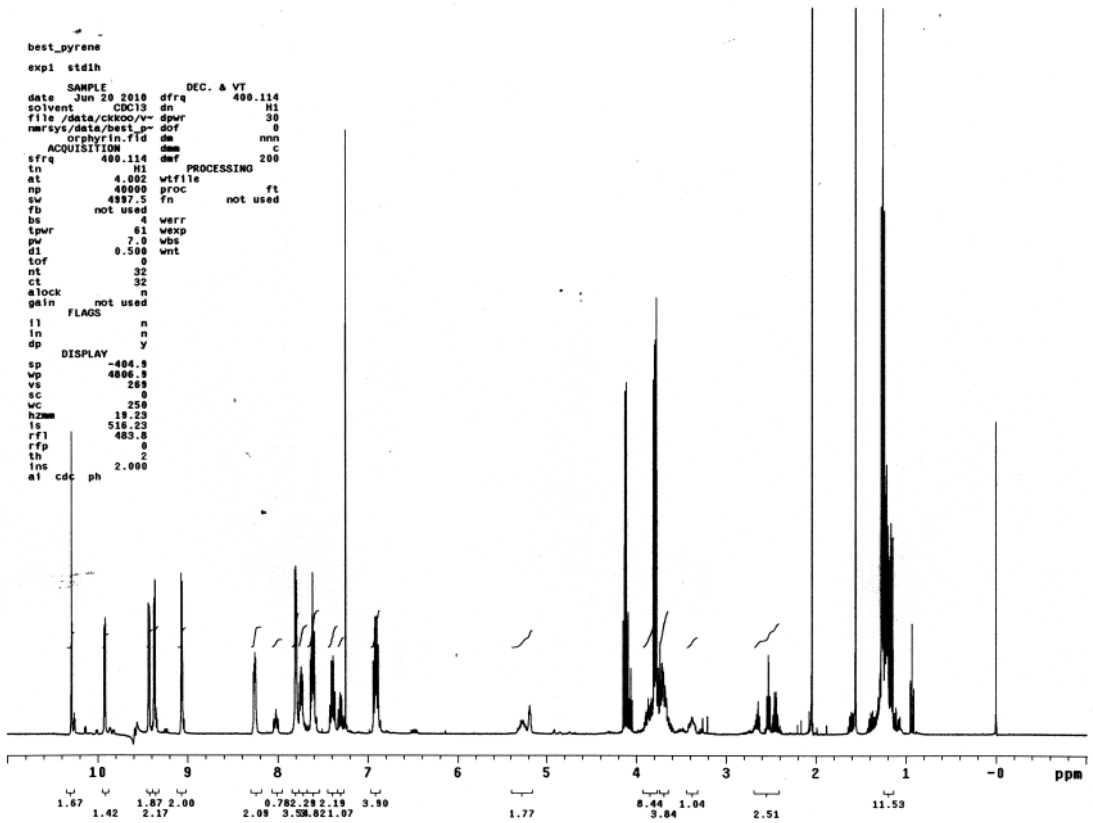


Figure S4. ^{1}H-NMR spectrum of **3** ($CDCl_3$).

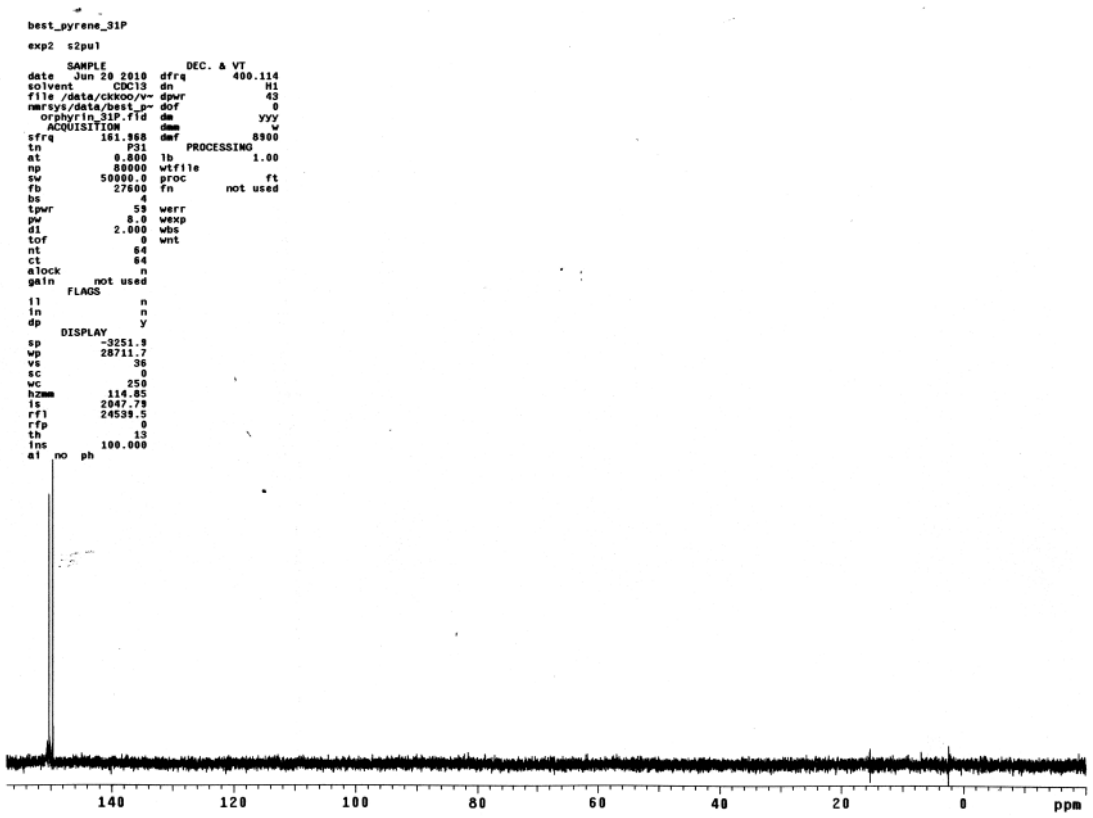


Figure S5. ^{31}P-NMR spectrum of **3** ($CDCl_3$).

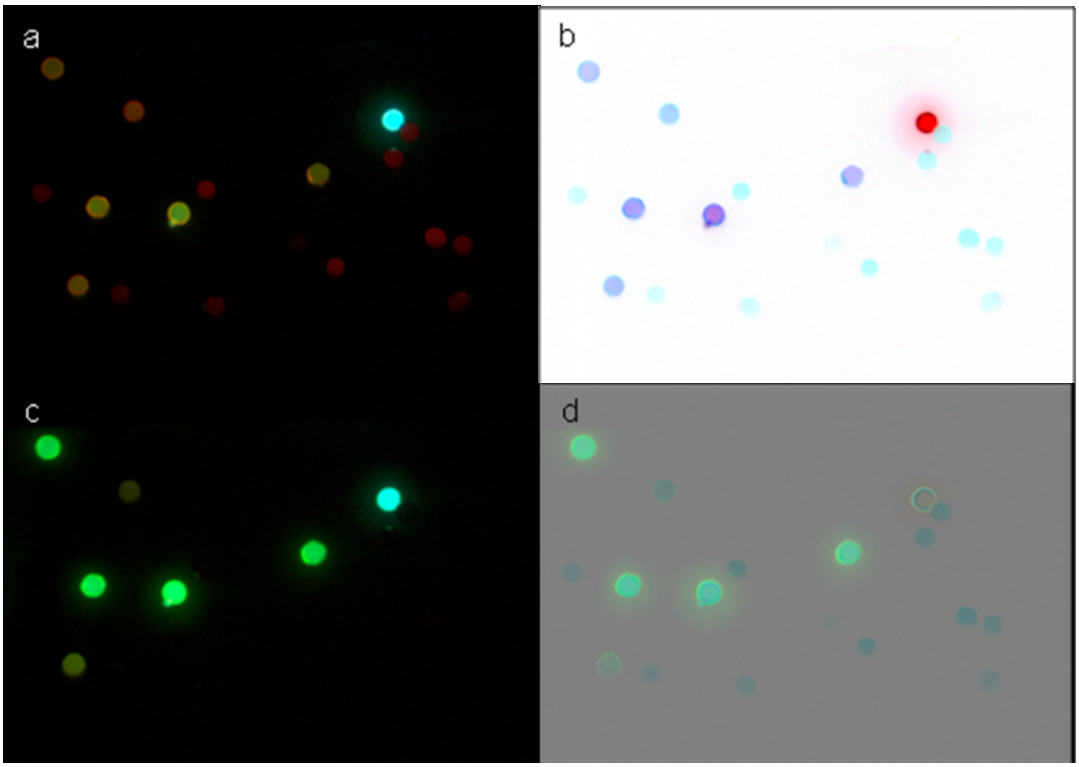


Figure S6. Examples of fluorescence images of the ODF library before and after exposure to HCl gas. (a) blank (before exposure); (b) inverted image of the blank; (c) after 5 min exposure to 10 ppt of HCl; (d) difference image, blending images (b) and (c) at 50% transparency. In the difference image (d), 50% gray represents no change after exposure; beads lighter than 50% gray indicate a lighting-up response, while darker represents a quenching response. Colors represent a combination of the original bead color and shifts in emission wavelength that occur upon exposure. Scale: beads are 130 microns in diameter.

Table S1. MS characterization of ODF dyes in this study.

	Sequence	Calculated Mass	Observed Mass
S1	**5'-DYYD-3'**	1601.4258 C84 H82 N2 O21 P4 Na	1601.4138 [a]
S2	**5'-DDPD-3'**	1847.4979 C94H97N7O21P4Zn	1847.7958 [a]
S3	**5'-DDDD-3'**	1621.5596 C84 H97 N4 O21 P4	1621.5583 [a]
S4	**5'-YYPE-3'**	1836.99 C98 H78 N4 O21 P4 Zn	1837.63 [b]
S5	**5'-YEYE-3'**	1639.45 C92 H74 O21 P4	1638.74 [b]
S6	**5'-DPYY-3'**	1804.3744 C94 H82 N5 O21 P4 Zn	1804.3689 [a]
S7	**5'-EDDY-3'**	1629.4595 C88 H85 N2 O21 P4	1629.4536 [a]
S8	**5'-EYPY-3'**	1836.99 C98 H78 N4 O21 P4 Zn	1836.70 [b]
S9	**5'-PPDD-3'**	2071.4207 C104 H95 N10 O21 P4 Zn2	2071.4243 [a]
S10	**5'-YEEE-3'**	1689.51 C96 H76 O21 P4	1689.64 [b]
S11	**5'-YDYY-3'**	1558.3860 C84 H76 N O21 P4	1558.3829 [a]
S12	**5'-DEDD-3'**	1650.5174 C88 H92 N3 O21 P4	1650.5106 [a]
S13	**5'-EDPY-3'**	1854.3901 C98 H84 N5 O21 P4 Zn	1854.3906 [a]
S14	**5'-DEDY-3'**	1629.4595 C88 H85 N2 O21 P4	1629.4575 [a]
S15	**5'-EDYP-3'**	1854.3901 C98 H84 N5 O21 P4 Zn	1854.3850 [a]

[a] HR-MS (negative mode, as $[(M-H)^-]$); [b] MALDI-TOF MS

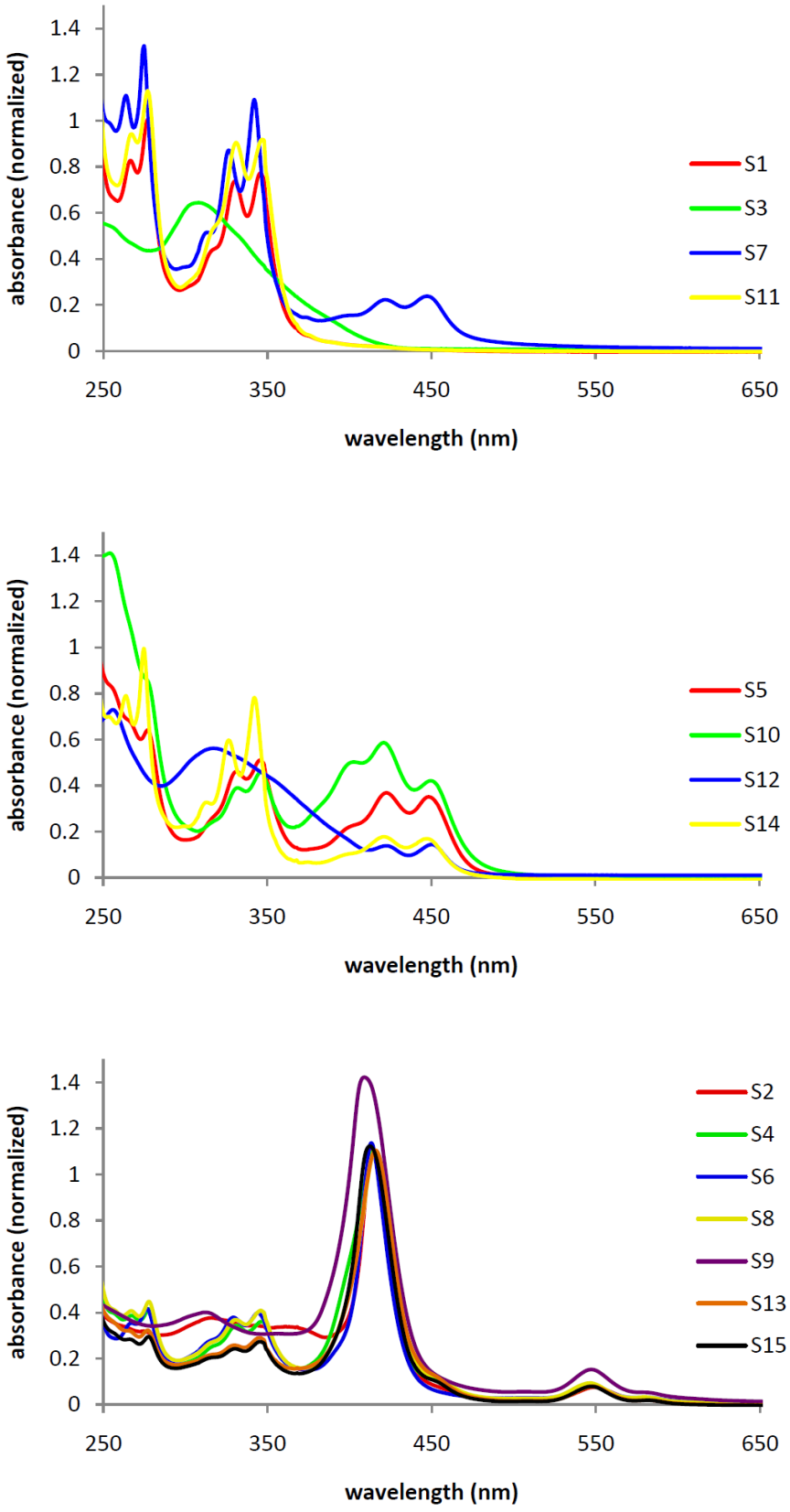


Figure S7. Normalized absorption spectra of the 15 ODFs from Table S1 in water.

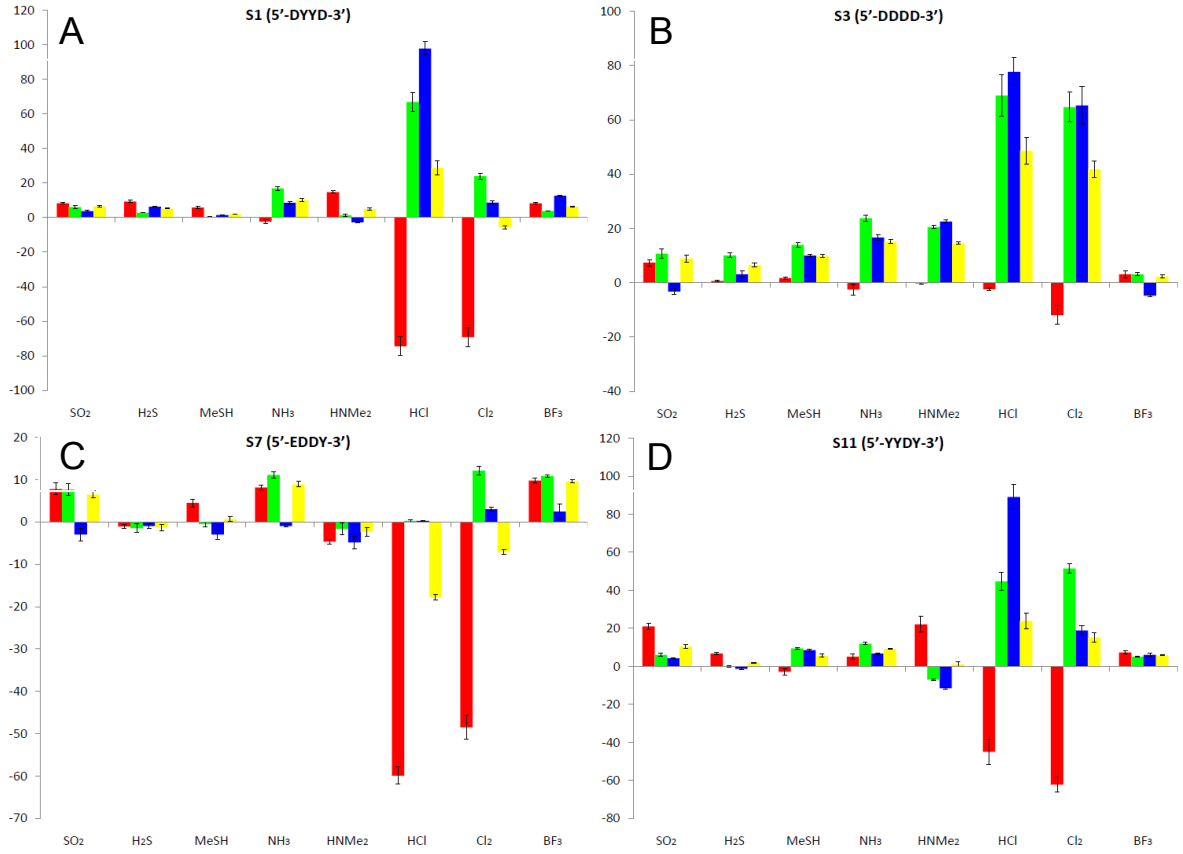


Figure S8. Quantitative color-change profiles of sensor sequences S**1** (A), S**3** (B), S**7** (C) and S**11** (D), showing spectral changes upon exposure to the eight selected toxic gases (1000 ppm). Error bars show variance in data averaged over 5 sensor beads.

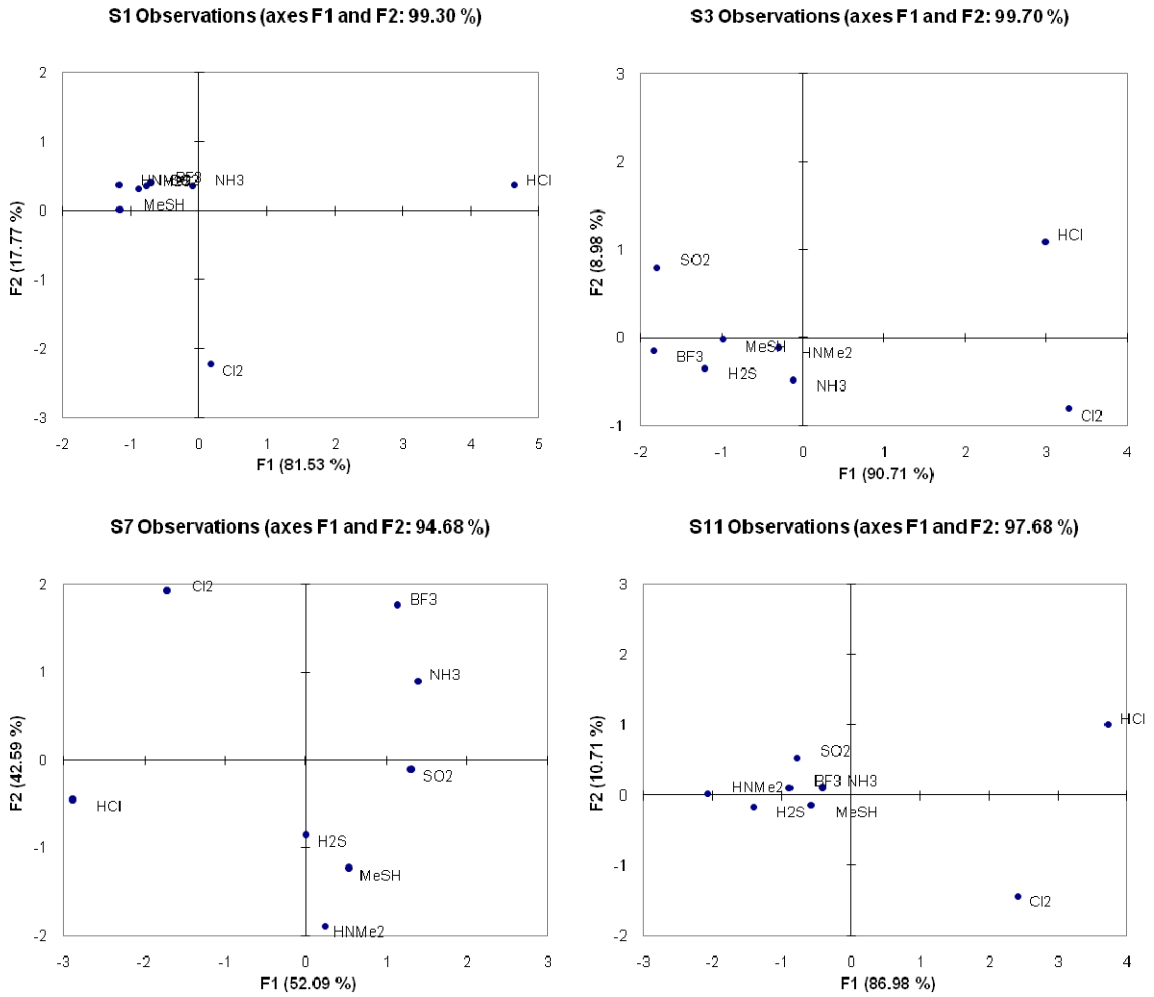


Figure S9. PCA scattering analysis of four of the the sensors' responses to the eight analytes, showing which analytes were the furthest and closest in responses. Note that these are 2-D representations of multidimensional data. Data are shown for **S1**, **S3**, **S7** and **S11**.

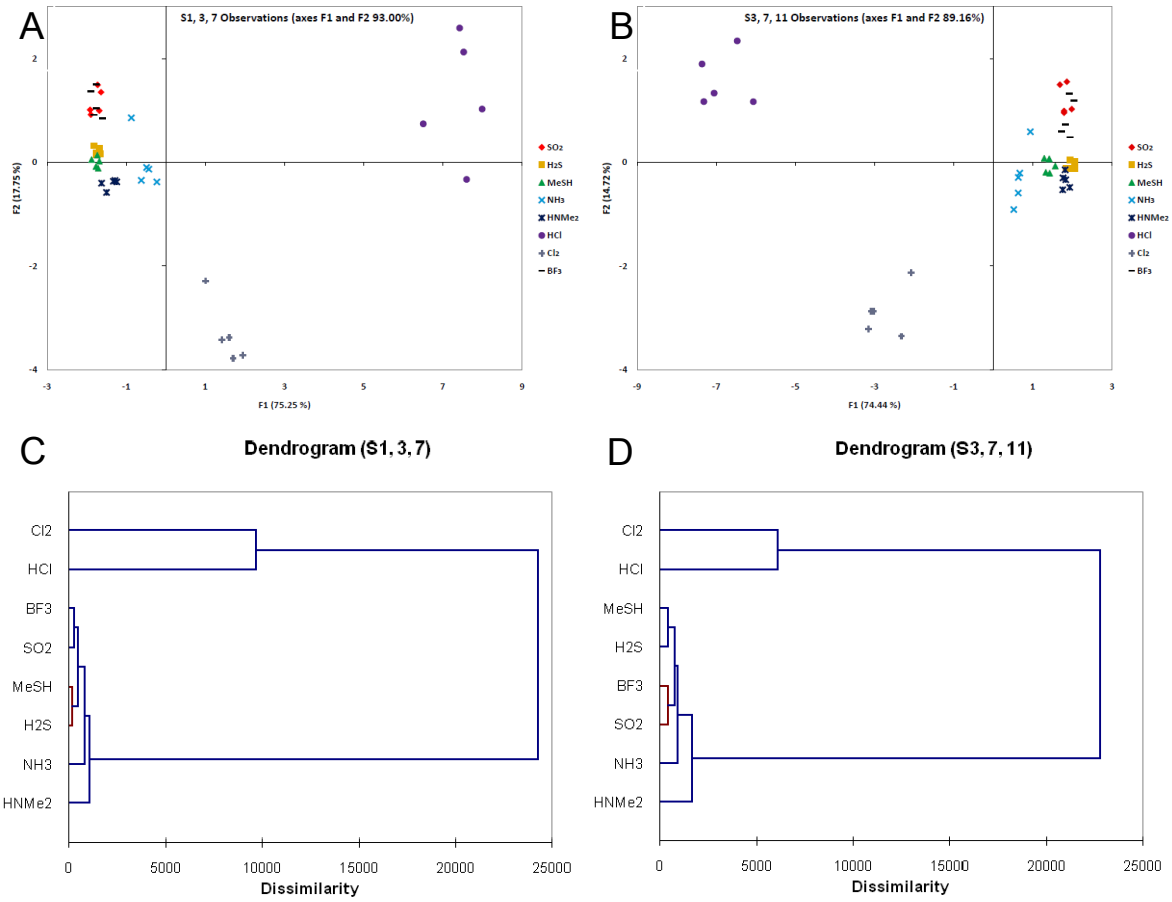


Figure S10. Strong responses with Cl_2 and HCl obscure the separation and analysis of the other 6 analytes. Data are shown for two three-sensor sets (**S1**, **S3**, **S7** and **S3**, **S7**, **S11**) for distinguishing the eight analytes. PCA scatter plots of the two major components (F1, F2) are shown in (A), (B); AHC dendrograms are in (C), (D).

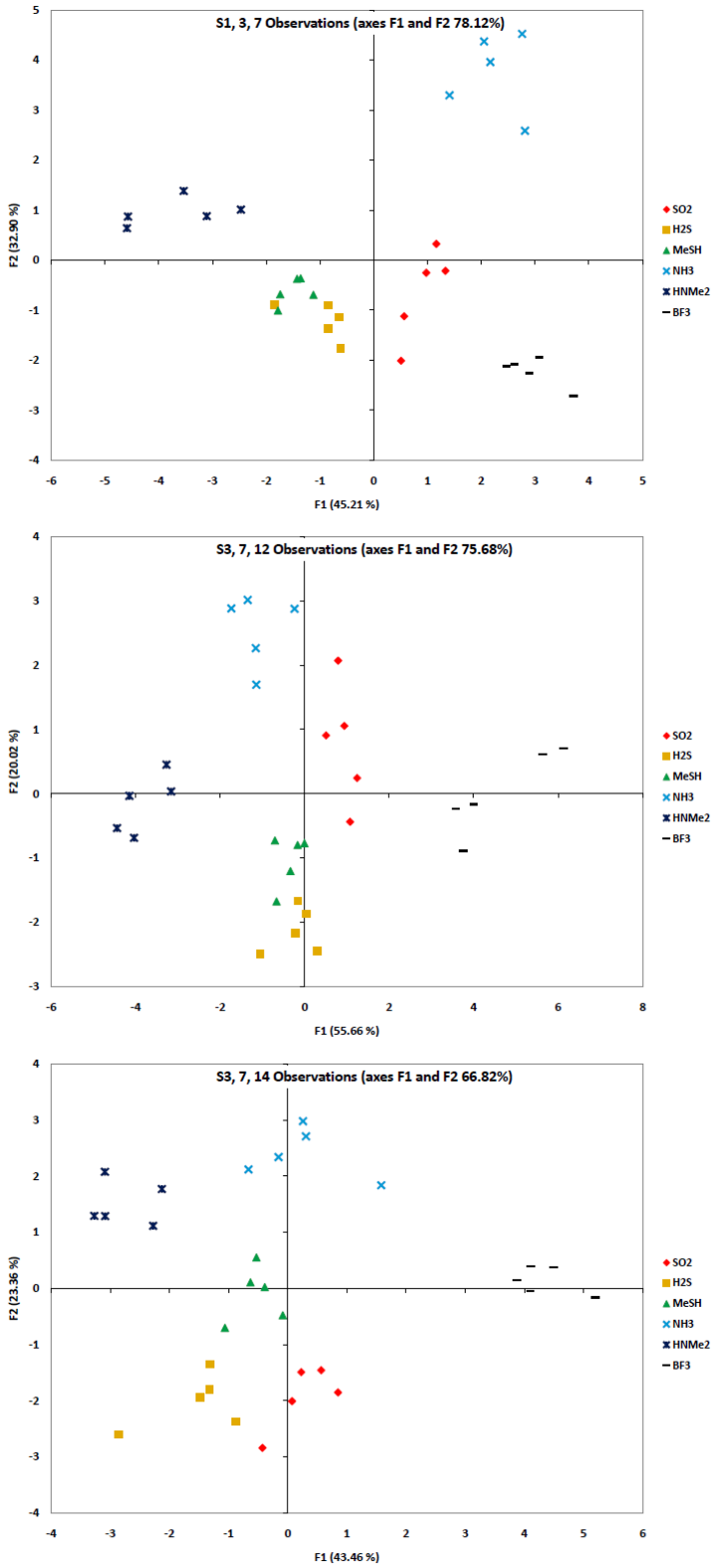


Figure S11. PCA scattering analysis of the reponses from three sensor groups (S**1**/S**3**/S**7**), (S**3**/S**7**/S**12**) and (S**3**/S**7**/S**14**) toward six analytes (Cl_2 and HCl omitted from the analysis). Note that these are 2-D plots of multidimensional data, and thus underestimate the actual separation of analytes.

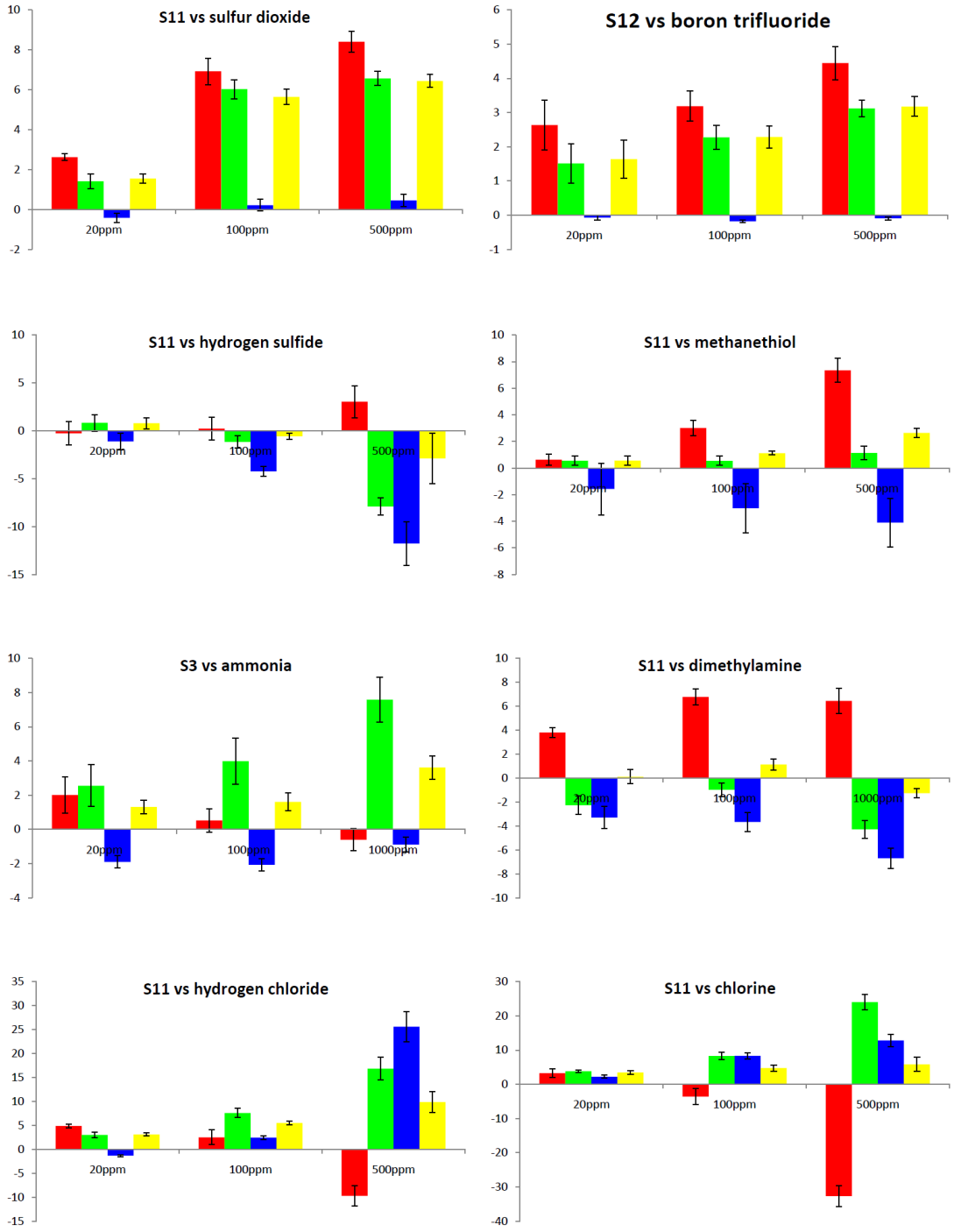


Figure S12. Responses of selected sensors upon exposure to varied concentrations of toxic gases in atmospheric air. Error bars show variance in data averaged over five sensor beads for each experiment.

Table S2. IDLH concentrations[6] of the selected toxic gases.

Name	**Molecular Formula**	**IDLH (ppm)**
Sulfur dioxide	SO_2	100
Hydrogen sulfide	H_2S	100
Methanethiol	MeSH	150
Hydrogen chloride	HCl	50
Chlorine	Cl_2	10
Boron trifluoride	BF_3	25
Ammonia	NH_3	300
Dimethylamine	$NHMe_2$	500

References

[1] (a) Ren, R. X.-F.; Chaudhuri, N. C.; Paris, P. L.; Rumney, S. IV; Kool, E. T. *J. Am. Chem. Soc.* **1996**, *118*, 7671-678.(b) Gao, J.; Strassler, C.; Tahmassebi, D.; Kool, E. T. *J. Am. Chem. Soc.* **2002**, *124*, 11590-11591.

[2] Bergstrom, D.E.; Zhang, P.; Zhou, J. *J. Chem. Soc. Perkin Trans I* **1994**, 3029

[3] Wang, Q.M.; Bruce, D.W. *Synlett* **1995**, 1267.

[4] Ohlmeyer, M. H. J.; Swanson, M.; Dillard, L. W.; Reader, J. C.; Asouline, G.; Kobayashi, R.; Wigler, M.; Still, W. C *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10922-0926.

[5] Gao, J.; Watanabe, S.; Kool, E. T. *J. Am. Chem. Soc.* **2004**, *126*, 12748-2749.

[6] http://www.cdc.gov/niosh/idlh/intridl4.html