Supporting Information

DNA as an Environmental Sensor: Detection and Identification of Pesticide Contaminants in Water With Fluorescent Nucleobases

Hyukin Kwon, Ke Min Chan, and Eric T. Kool*

Department of Chemistry, Stanford University, Stanford, CA 94305 (USA); e-mail: kool@stanford.edu

Table of Contents

Experimental Methods (p. 2)
Table S1. List of pesticides used in this study. (p. 6)
Table S2. MALDI-TOF mass spectrometry data of the resynthesized ODF sequences. (p. 7)
Figure S1. Absorption spectra of resynthesized ODF sequences. (p. 8)
Figure S2. Fluorescence emission spectra of resynthesized ODF sequences. (p. 9)
Figure S3. Representative bead images showing fluorescence of the ODFs before and after metal incorporation (p. 10)
Figure S4. Quantitative fluorescence responses (as measured by ΔRGB from microscopy images) of the twenty-eight ODF sensors versus the fifteen pesticides. (p. 14)
Figure S5. Comparison of responses from isomeric chemosensors, with the same monomers in a different sequence (p. 15)
Figure S6. Comparison of responses of identical chemosensors with different metals. (p. 16)
Figure S7. Data used for estimating concentrations of paraquat in unknown samples. (p. 17)
Figure S8. Data used for estimating glyphosate concentrations in unknown samples. (p. 18)
Figure S9. Data used for attempted estimation of dimethoate concentrations in unknowns. (p. 19)
Figure S10. Blind identification of two unknown pesticides in lake water. (p. 20)
References (p. 21)

Experimental Methods

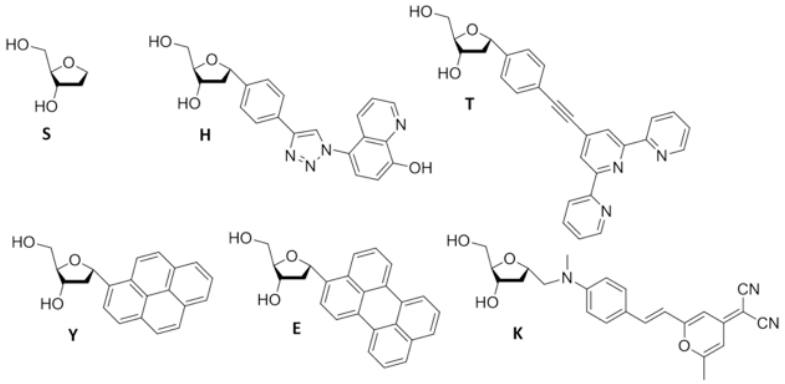
Chemicals and reagents

Anhydrous solvents were purchased from Acros Organics and used without further purification. Chemicals for oligodeoxyfluoroside (ODF) monomer synthesis, anion sodium salts, and metal nitrates were purchased from Sigma-Aldrich unless otherwise noted. All anion solutions were prepared within hours of use. Chemicals used for solid-phase synthesis of ODF were purchased from Glen Research, including spacer (S) phosphoramidites, 3'-phosphate CPG, synthesizer reagent solutions and deprotection reagent. All chemical reactions were performed under argon gas unless otherwise noted. Silica gel (60 Å, 200-425 mesh) was used for flash column chromatography.

General instrumentation.

^{1}H, ^{13}C, and ^{31}P NMR spectra were recorded using Varian Innova 400 MHz instrument unless otherwise noted. Internal signal from NMR solvents ($CDCl_3$ or DMSO-d_6) were used as references. Chemical shifts are reported as ppm, and multiplicity patterns are abbreviated as the following: singlet (s), doublet (d), triplet (t), and multiplet (m). Mass spectra were obtained using ESI or MALDI-TOF at Stanford University Mass Spectrometry Facility and Stanford Protein and Nucleic Acid Facility, respectively. Gas chromatography was performed using Shimazu GC17A instrument (EC detector). HPLC was performed using Shimazu LC-20AD (SPD-M20A diode array detector) and reverse phase C5 column (Phenomenex Jupiter). Absorption spectra were obtained using Varian Cary 100 Bio UV-Vis Spectrophotometer. Steady-state fluorescence emission spectra were measured on Jobin Yvon-Spex Fluorolog 3 spectrometer.

ODF library construction.



ODF fluorescent deoxyriboside monomer H was synthesized and derivatized as the 5'-DMT, 3'-phosphoramidite derivative as described.[1] Monomers Y and E were prepared according to the literature methods.[2] Compound K was synthesized in 14 steps following the published methods.[3] The phosphoramidite derivative of monomer T was prepared followed the published procedure.[4] Spacer monomer S phosphoramidite was purchased from Glen Research. The library was assembled on amine-functionalized polyethylene glycol-polystyrene beads (PS beads, 130 μm, NovaSyn TG amino resin) as previously reported[4] to yield 1296 unique sequences of tetramers. Binary chemical tags[5] were installed during the library synthesis and later cleaved for gas chromatographic sequence identification (see below). ODFs were deprotected using 50 mM potassium carbonate in methanol, washed with EDTA in

dimethylformamide (DMF), water, acetonitrile, and lastly with dichloromethane (DCM), and dried using argon stream.

Library screening.

Approximately fifty beads at a time from the library were pre-equilibrated with either $Y(NO_3)_3 \cdot 4\ H_2O$ or $Zn(NO_3)_2 \cdot 6H_2O$. In some cases, beads that were not exposed to any metal were used for screening. The beads were shaken in 1 mL of 25 mM metal solution in acetonitrile for 30 minutes and then thoroughly washed with water and acetonitrile. They were placed on a small square removable double-sided tape (5 mm, 3M Scotch) attached to a Petri dish (35 mm diameter, Falcon). 25 μL of 1 mM Tris-HCl buffer (pH 8) was added and allowed to equilibrate for one hour at room temperature protected from light. A digital image was captured under epifluorescence microscope (Nikon Eclipse 80i, equipped with Nikon Plan Fluor 4x/0.13 objective, ND8 filter, and QIClick digital CCD camera) using λ_{ex} = 340-380 nm and λ_{em} > 420 nm filters. The exposure times were set constant during screening (80 ms for each RGB channel, gain 4x, 24-bit image). Next, anion (500 μM) in the same buffer (25 μL) was added to above and incubated for 30 minutes at room temperature. A second fluorescence image was captured using the same microscope setting. Comparing the "before" and "after" picture visually, noticeable and desirable fluorescence color changes were noted and these beads were isolated. They were placed in a sealed capillary tube, and the chemical tags were released using 3 μL of CAN solution (0.5 M ceric ammonium nitrate in 1:1 water:acetonitrile) and 3 μL of decane. The capillary tubes were sonicated for three hours, centrifuged briefly, and the organic layer was derivatized with *N*,*O*-bis(trimethylsilyl)acetamide and analyzed with gas chromatography to decode the sequence.[5]

Preparing pesticide solutions

The sixteen pesticides were purchased from Sigma-Aldrich. Pesticide solutions (100 μM) for testing were prepared from 1 mM stock in 1 mM Tris-HCl buffer (pH 8.0). All pesticide solutions were prepared within hours of use. For pesticides with low solubilities (atrazine, cyanazine, diuron, trifluralin, carbaryl, and malathion), 100 μM solutions were prepared directly (100 mL) and heated to 50 °C for ten minutes while stirring, and then sonicated for another ten minutes to ensure dissolution.

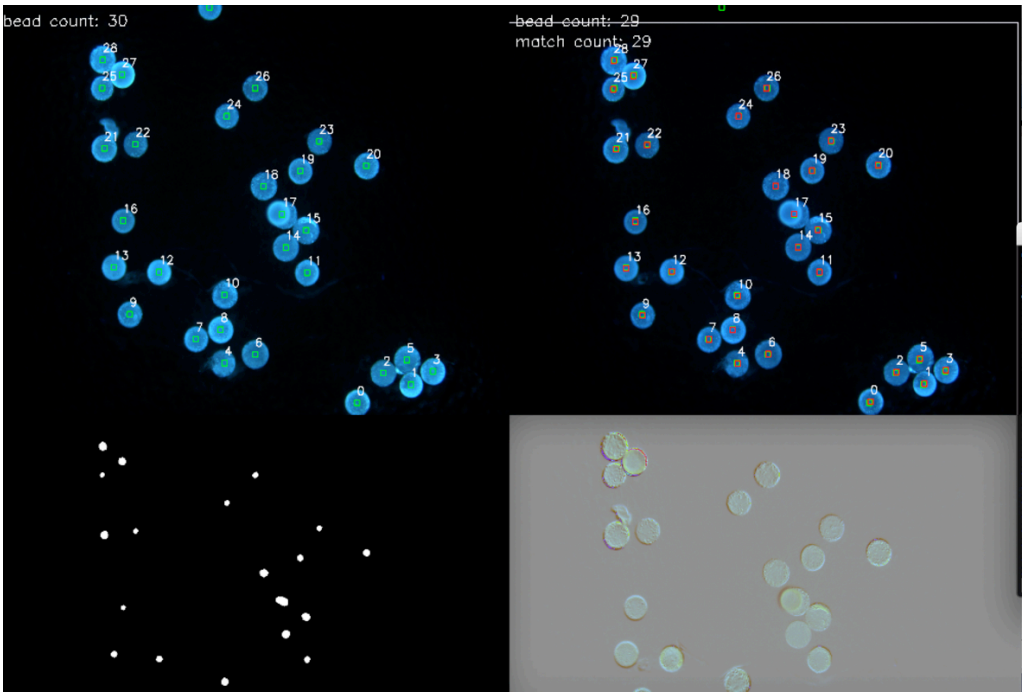
Resynthesis and characterization of screened ODF sensors

The screened ODF sequences were resynthesized on ABI 394 DNA synthesizer using standard phosphoramidite oligonucleotide synthesis (coupling time of 15 min). Both 3'-phosphate CPG (1 μmol, Glen Research) and 10 mg dimethoxytrityl (DMT)-functionalized PEG-PS beads (0.29 mmol/g) were added to DNA synthesis columns to allow simultaneous synthesis of both cleavable and solid-phase ODF sequences. The final yields after tetramer synthesis was above 50 % for all 29 sequences. The two solid supports were separated after synthesis in DCM, and each was deprotected with 50 mM potassium carbonate in methanol to afford cleaved ODF sequence in solution and ready-to-use ODF on PEG-PS beads. The solid-phase ODFs were washed with water and acetonitrile and dried with argon before use. Cleaved ODFs were filtered, dried, and purified by HPLC using C5 reverse-phase column and 50 mM triethylammonium acetate and acetonitrile as mobile phase. The purified sequences were redissolved in 0.5 mL water and characterized by MALDI-TOF mass spectrometry, and its optical properties (absorption and fluorescence emission spectra were measured in phosphate-buffered saline (PBS).

Resynthesized sensors reaction with pesticides and data acquisition

Resynthesized ODF sensors on PEG-PS beads were pre-equilibrated with one of the two metals (where applicable) and placed in a Petri dish (35mm diameter) using a small square piece (5 mm) of removable double-sided tape. They were incubated in 3 mL of 1 mM Tris-HCl buffer (pH 8) for one hour at room temperature protected from light, and the "before" image was taken under epifluorescence microscope using λ_{ex} = 340-380 nm and λ_{em} > 420 nm filters. The exposure times were constant for each sequence throughout the experiments and ranged from 50 ms to 200 ms (for all RGB channels, gain 4x) to avoid over-exposure. The solution was then replaced with 3 mL of the same buffer containing the analyte of interest at 100 μM, and the "after" image was captured after 30 minutes. Four beads were randomly chosen and a 15-by-15 pixel box was placed at the center of each bead. Mean red (R), green (G), and blue (B) channel values over the pixel selection were extracted and the difference values (ΔR, ΔG, and ΔB with theoretical range from -255 to +255) were obtained, with error indicating standard deviation from the mean of the four beads.

Custom machine vision software (Python script based on the OpenCV framework) was written in-house to facilitate the rapid data extraction of ODF bead color changes in response to presence of pesticides (see screenshot below). The bead image analysis software will be made available upon request.



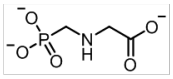
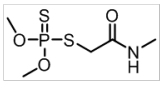
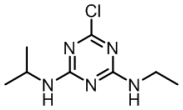
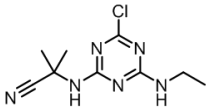
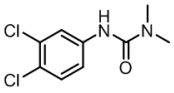
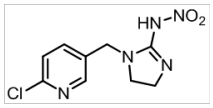
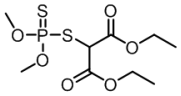
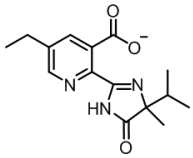
Upper left quadrant of screenshot: ODF sequence YYYY on polystyrene beads *before* analyte exposure. Average fluorescence response of each bead was sampled from a 15x15 pixel grid at the center of the bead (green boxes) in automated fashion. *Upper right quadrant*: ODF beads after analyte exposure. The ECC Image Alignment Algorithm from the OpenCV framework was used to estimate the geometric warp between the *before* and *after* images, and an affine transformation was applied to the *after* image in order to align the two images. The identities of the ODF beads were assigned based on the spatial position of the bead centers. Red boxes indicate the 15x15 pixel grids at the repositioned bead centers from which the

RGB values were sampled. *Lower left quadrant*: visualization of the positions of the bead centers. Image thresholding was performed to keep the bright bead regions of each image, then edge regions closest to zero pixels were eroded. The centroid of each contour was taken as the bead center. *Lower right quadrant*: visualization of color difference between the *before* and *aft*er images. Subtraction of the pixel intensities of the *before* from the *after* image, followed by normalization of mean pixel intensity.

Statistical methods

ΔR, ΔG, and ΔB from each bead were used for discriminant analysis (DA) and agglomerative hierarchical clustering (AHC). Addinsoft XLSTAT was used to generate both analyses. For DA, ellipses around the centroid represent 95% confidence limit. For AHC, dissimilarity was analyzed on Euclidean distance using Ward's agglomeration method.

Table S1. List of pesticides used in this study. The full name, abbreviation, and estimated maximum solubility in water are shown.

Name and abbreviation	Structure	Solubility in water (μM, approximate)
Glyphosate GLY (herbicide)		60 000
Dimethoate DIM (insecticide)		110 000
Atrazine ATR (herbicide)		320
Cyanazine CYA (herbicide)		830
DCMU (Diuron) DIU (herbicide)		180
Imidacloprid IMI (insecticide)		1 990
Malathion MAL (insecticide)		440
Imazethapyr IMA (herbicide)		480 000

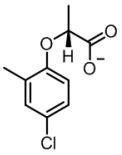
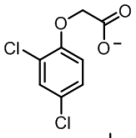
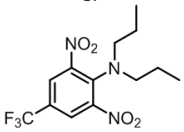
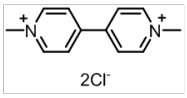
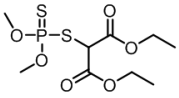
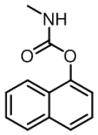
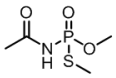
Name and abbreviation	Structure	Solubility in water (μM, approximate)
Mecoprop-p MEC (herbicide)		4 200
2,4-D 24D (herbicide)		4 100
Trifluralin TRI (herbicide)		100
Paraquat PAR (herbicide)		High
Metolachlor MET (herbicide)		1 900
Carbaryl CAR (insecticide)		50
Acephate ACE (insecticide)		4 300 000

Table S2. MALDI-TOF mass spectrometry data of the resynthesized ODF chemosensor sequences.

Sequence	Expected Mass	Mass Found
EHHS	1560.35	1560.91
EKYS	1489.35	1488.61
ESKY	1489.35	1489.65
HEKS	1575.37	1574.92
HHHH	1882.45	1883.68
HHKH	1897.48	1897.77
HSSH	1310.27	1310.35
HSSS	1024.18	1024.14
HSSY	1224.24	1224.32
HSYY	1424.30	1424.68
KHYS	1525.36	1525.24
KTYY	1770.44	1771.67
SHEH	1560.35	1561.56
SHYS	1224.24	1224.16
SSSY	938.15	938.41
SYYE	1388.29	1389.16
SYYH	1424.30	1424.92
TSSS	1069.20	1069.91
TTYE	1850.45	1851.78
TYHS	1555.35	1556.04
TYSY	1469.32	1470.46
YEYH	1674.38	1676.39
YHSY	1424.30	1424.78
YSKY	1439.33	1438.84
YSYY	1338.29	1338.38

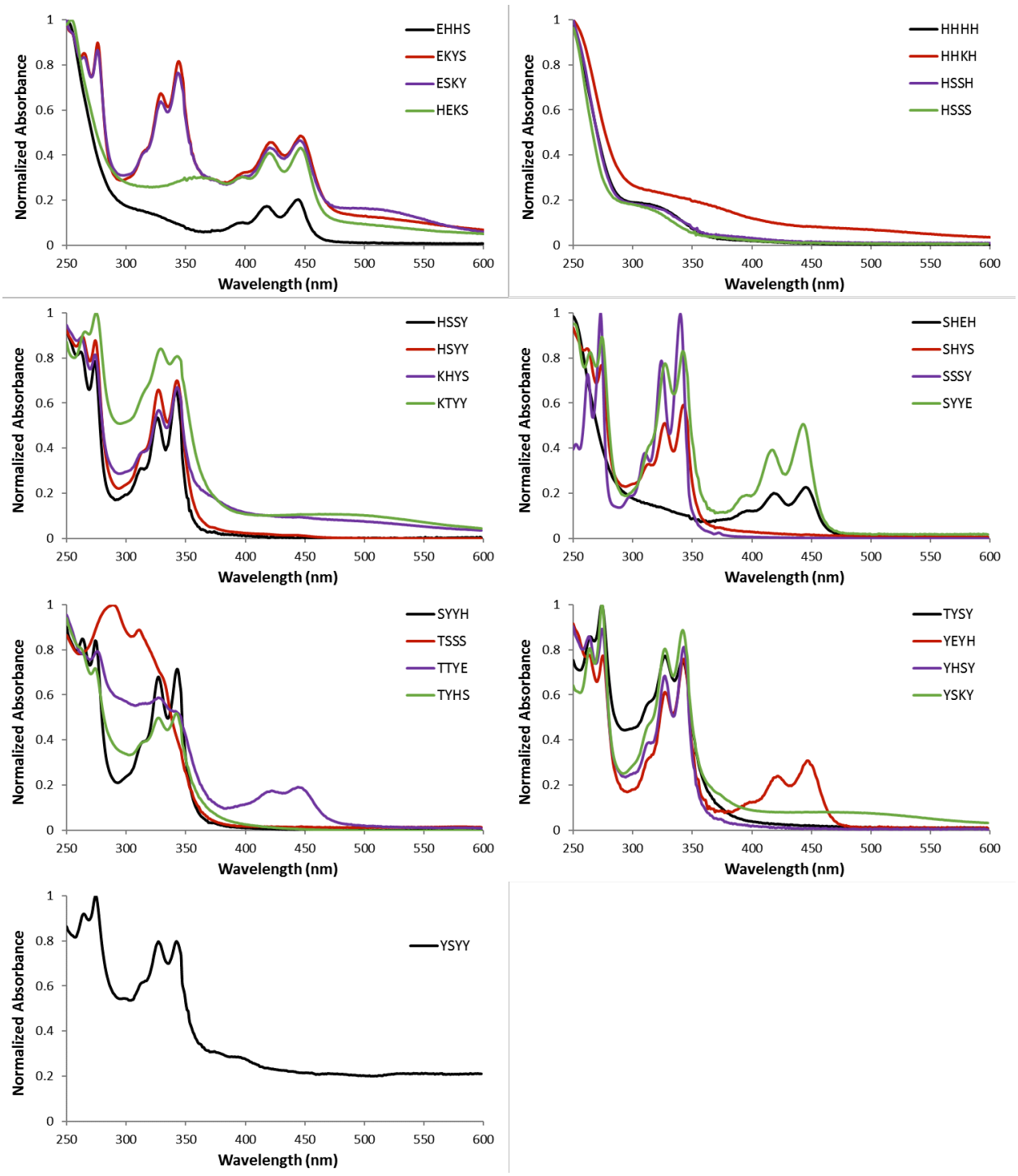


Figure S1. Absorption spectra of resynthesized ODF chemosensor sequences (20 uM) in phosphate-buffered saline.

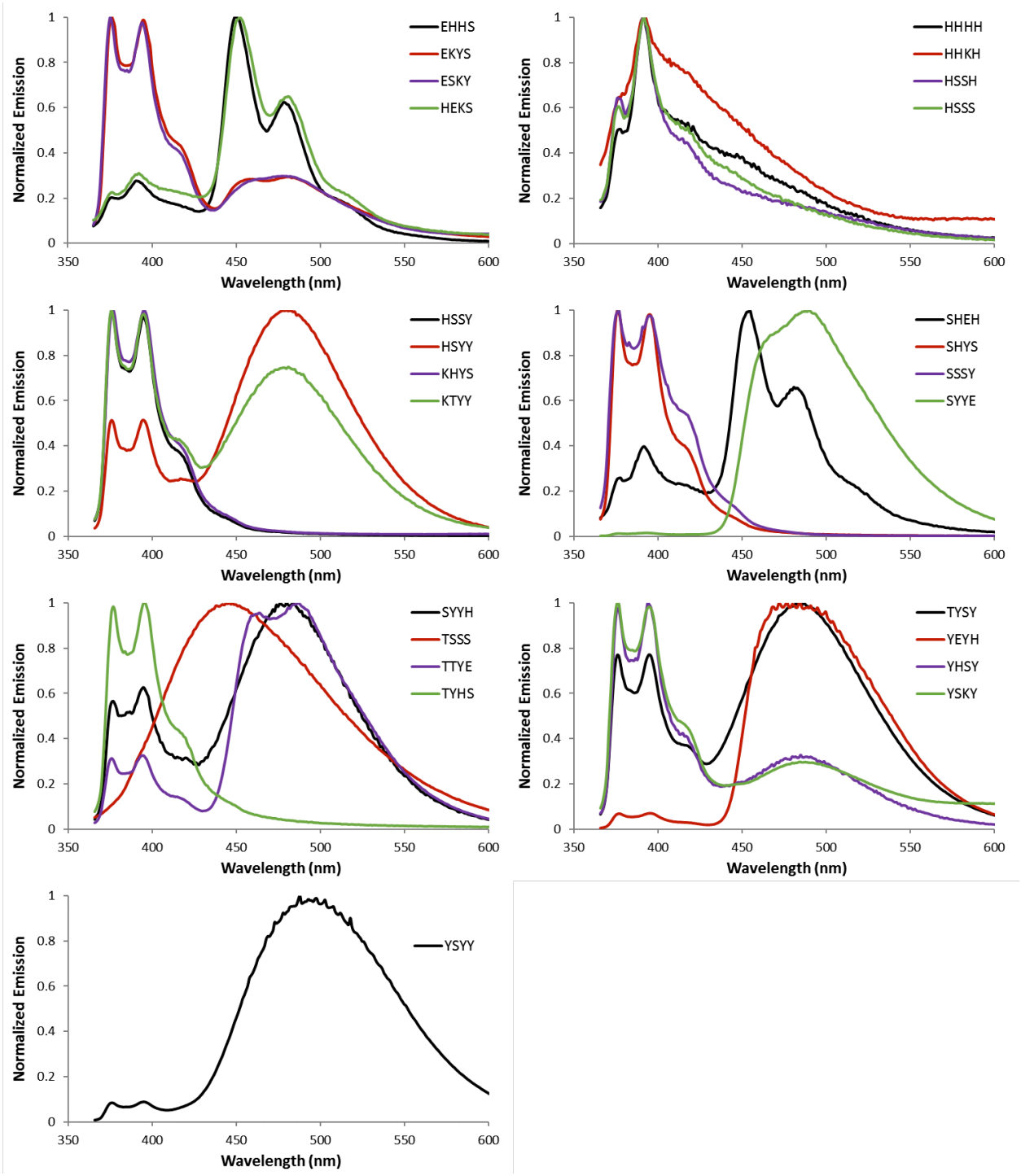
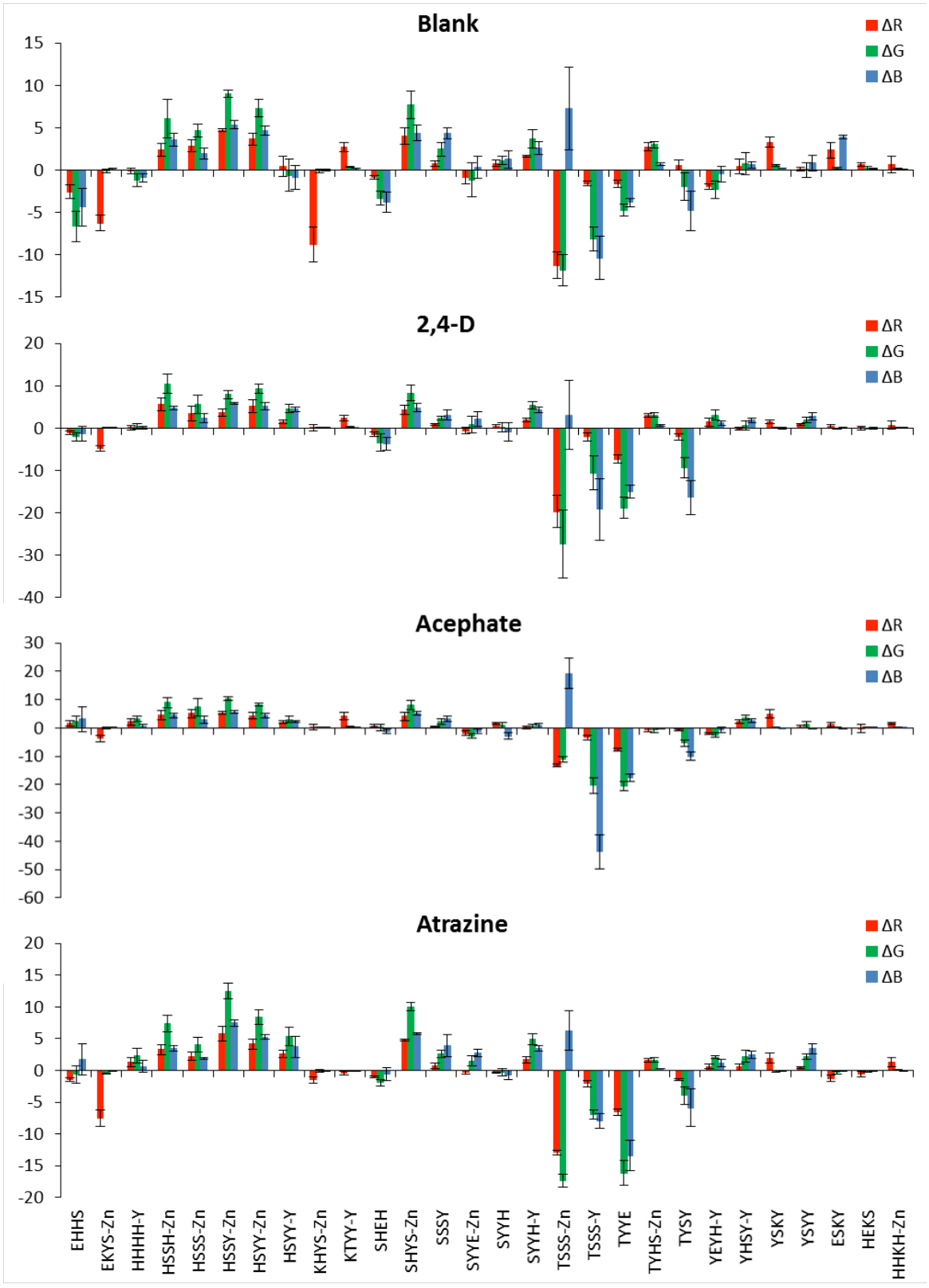
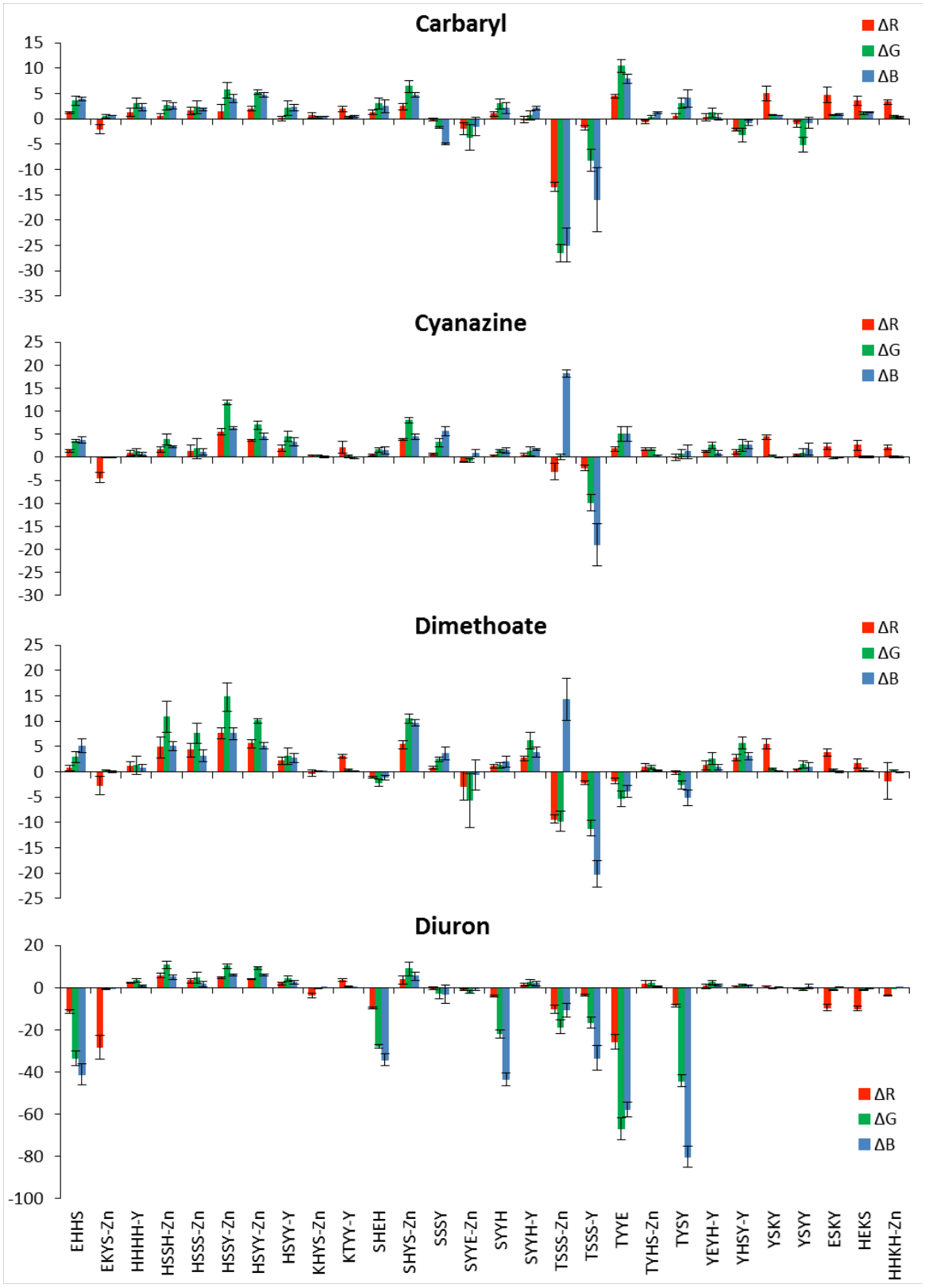
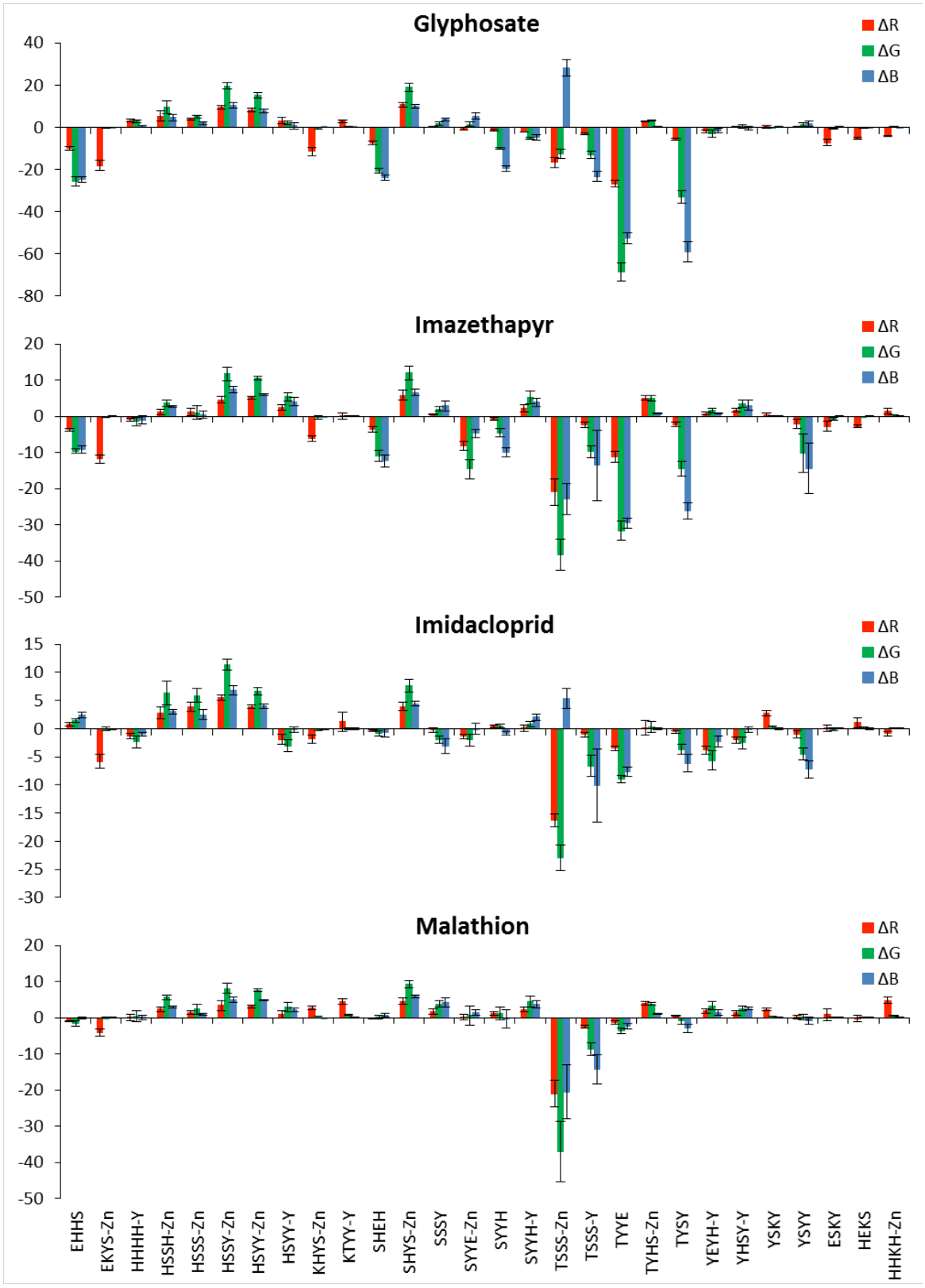


Figure S2. Fluorescence emission spectra of resynthesized ODF chemosensor sequences (20 μM) in PBS buffer (λ_{ex} = 345 nm, λ_{em} > 365 nm).

Figure S3. Representative bead images showing fluorescence of the ODFs before and after metal incorporation (where applicable). Sample images of twenty-eight resynthesized ODF tetramer sequences on PEG-PS beads were captured using epifluorescence microscope (λ_{ex} = 340-380 nm, λ_{em} > 420 nm). The beads on the left side of each column are ODFs without added metal, and those on the right are after exposure to Y^{III} or Zn^{II} metal nitrate salts. All beads were imaged after incubation in 1 mM Tris-HCl buffer (pH 8.0) for one hour at room temperature. Each bead represents the average color of the beads in each sample.


Blank
ΔR
ΔG
ΔB
15
10
5
0
-5
-10
-15
2,4-D
ΔR
ΔG
ΔB
20
10
0
-10
-20
-30
-40
Acephate
ΔR
ΔG
ΔB
30
20
10
0
-10
-20
-30
-40
-50
-60
Atrazine
ΔR
ΔG
ΔB
20
15
10
5
0
-5
-10
-15
-20
EHHS
EKYS-Zn
HHHH-Y
HSSH-Zn
HSSS-Zn
HSSY-Zn
HSYY-Zn
HSYY-Y
KHYS-Zn
KTYY-Y
SHEH
SHYS-Zn
SSSY
SYYE-Zn
SYYH
SYYH-Y
TSSS-Zn
TSSS-Y
TYYE
TYHS-Zn
TYSY
YEYH-Y
YHSY-Y
YSKY
YSYY
ESKY
HEKS
HHKH-Zn


Carbaryl
ΔR
ΔG
ΔB
15
10
5
0
-5
-10
-15
-20
-25
-30
-35
Cyanazine
ΔR
ΔG
ΔB
25
20
15
10
5
0
-5
-10
-15
-20
-25
-30
Dimethoate
ΔR
ΔG
ΔB
25
20
15
10
5
0
-5
-10
-15
-20
-25
Diuron
20
0
-20
-40
-60
-80
-100
ΔR
ΔG
ΔB
EHHS
EKYS-Zn
HHHH-Y
HSSH-Zn
HSSS-Zn
HSSY-Zn
HSYY-Zn
HSYY-Y
KHYS-Zn
KTYY-Y
SHEH
SHYS-Zn
SSSY
SYYE-Zn
SYYH
SYYH-Y
TSSS-Zn
TSSS-Y
TYYE
TYHS-Zn
TYSY
YEYH-Y
YHSY-Y
YSKY
YSYY
ESKY
HEKS
HHKH-Zn


Glyphosate
ΔR
ΔG
ΔB
40
20
0
-20
-40
-60
-80
Imazethapyr
ΔR
ΔG
ΔB
20
10
0
-10
-20
-30
-40
-50
Imidacloprid
ΔR
ΔG
ΔB
15
10
5
0
-5
-10
-15
-20
-25
-30
Malathion
ΔR
ΔG
ΔB
20
10
0
-10
-20
-30
-40
-50
EHHS
EKYS-Zn
HHHH-Y
HSSH-Zn
HSSS-Zn
HSSY-Zn
HSYY-Zn
HSYY-Y
KHYS-Zn
KTYY-Y
SHEH
SHYS-Zn
SSSY
SYYE-Zn
SYYH
SYYH-Y
TSSS-Zn
TSSS-Y
TYYE
TYHS-Zn
TYSY
YEYH-Y
YHSY-Y
YSKY
YSYY
ESKY
HEKS
HHKH-Zn

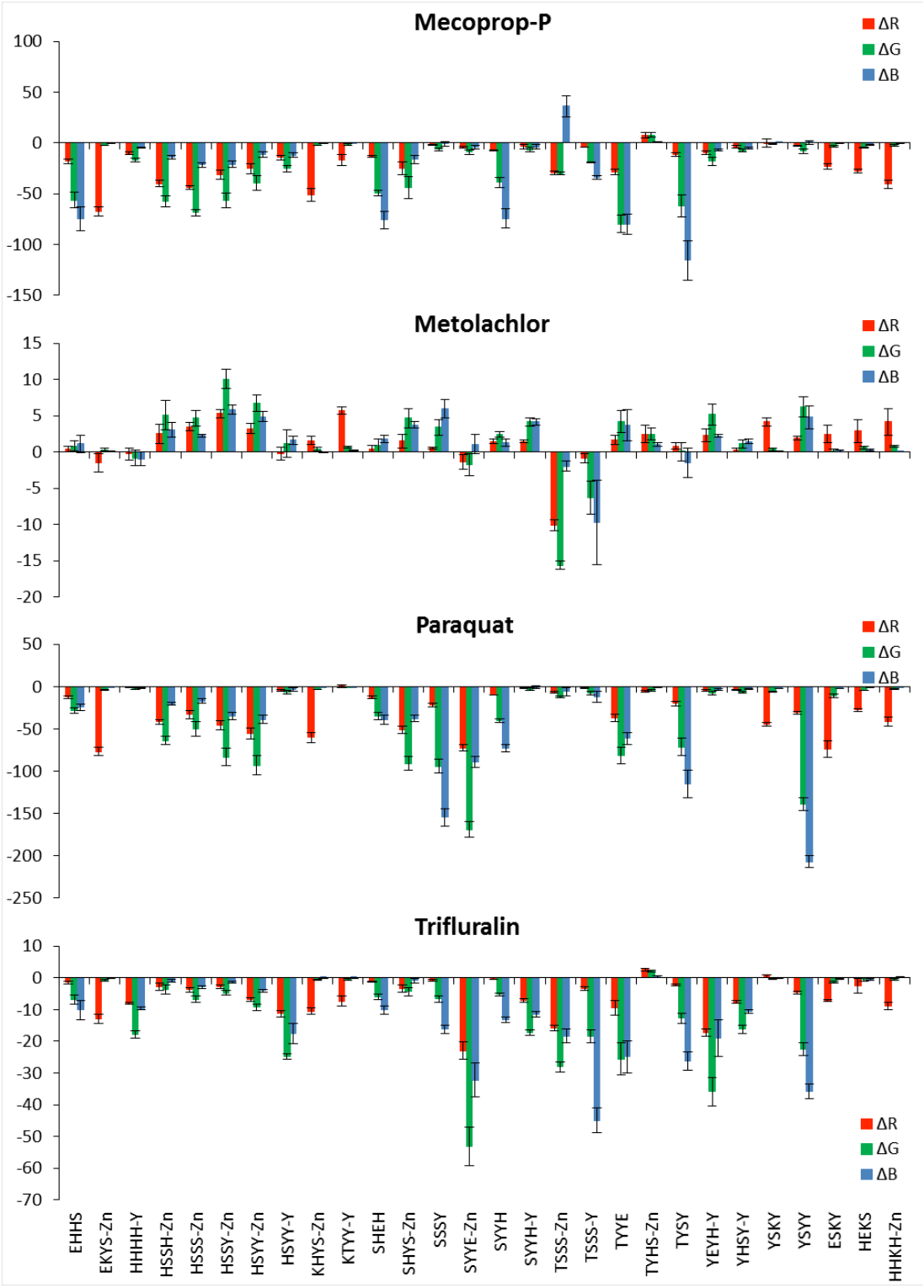


Figure S4. Quantitative fluorescence responses (as measured by ΔRGB from microscopy images) of the twenty-eight ODF sensors versus the fifteen pesticides at 100 uM after 30 min equilibration in Tris-buffered water (pH 8.0). Red, green, blue, and grey bars represent changes in R, G, B, and L (on a scale of ±255), respectively. The error bars represent standard deviation from four measurements.

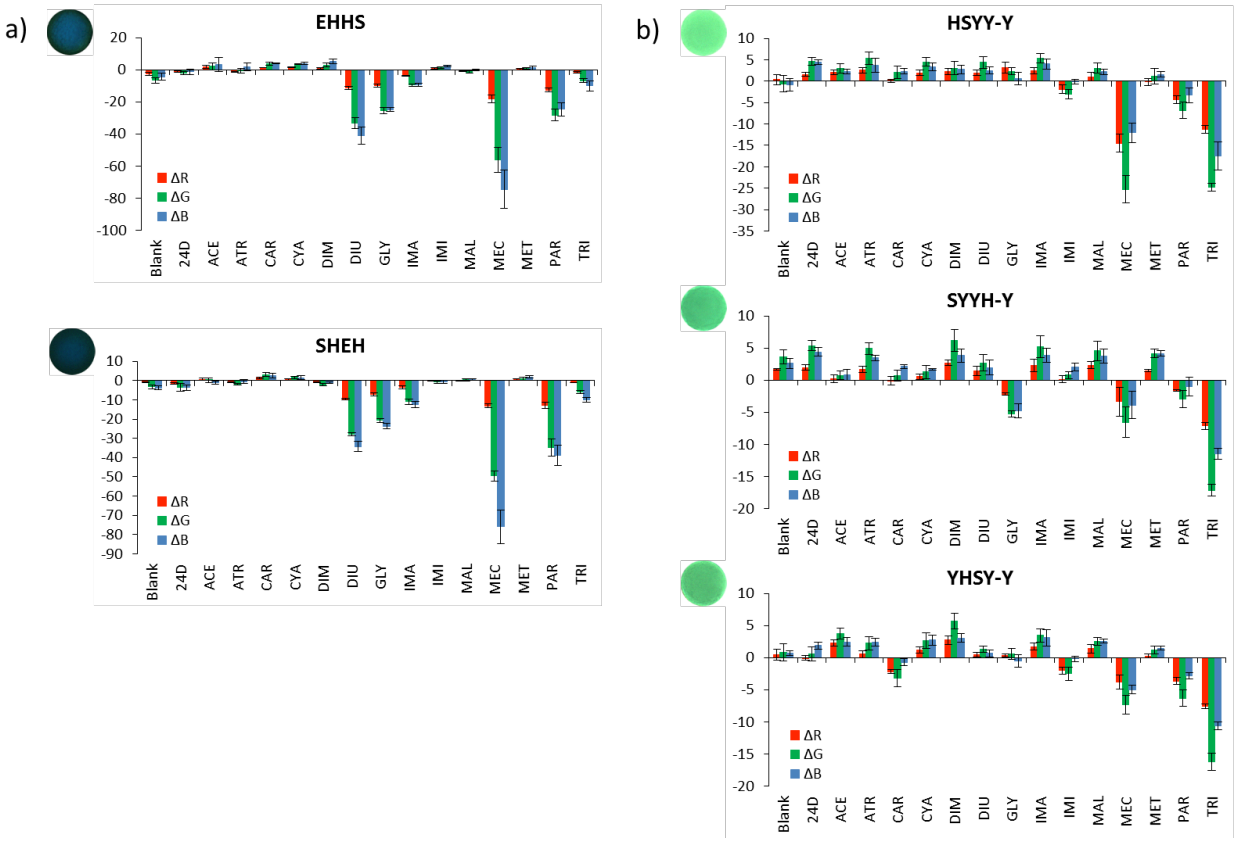


Figure S5. Comparison of responses from isomeric chemosensors, with the same monomers in a different sequence. Responses from two sets of isomers, (a) and (b), are shown. A sample bead image for each chemosensor sequence is also included. This data was collected from the data set represented in Figure S4.

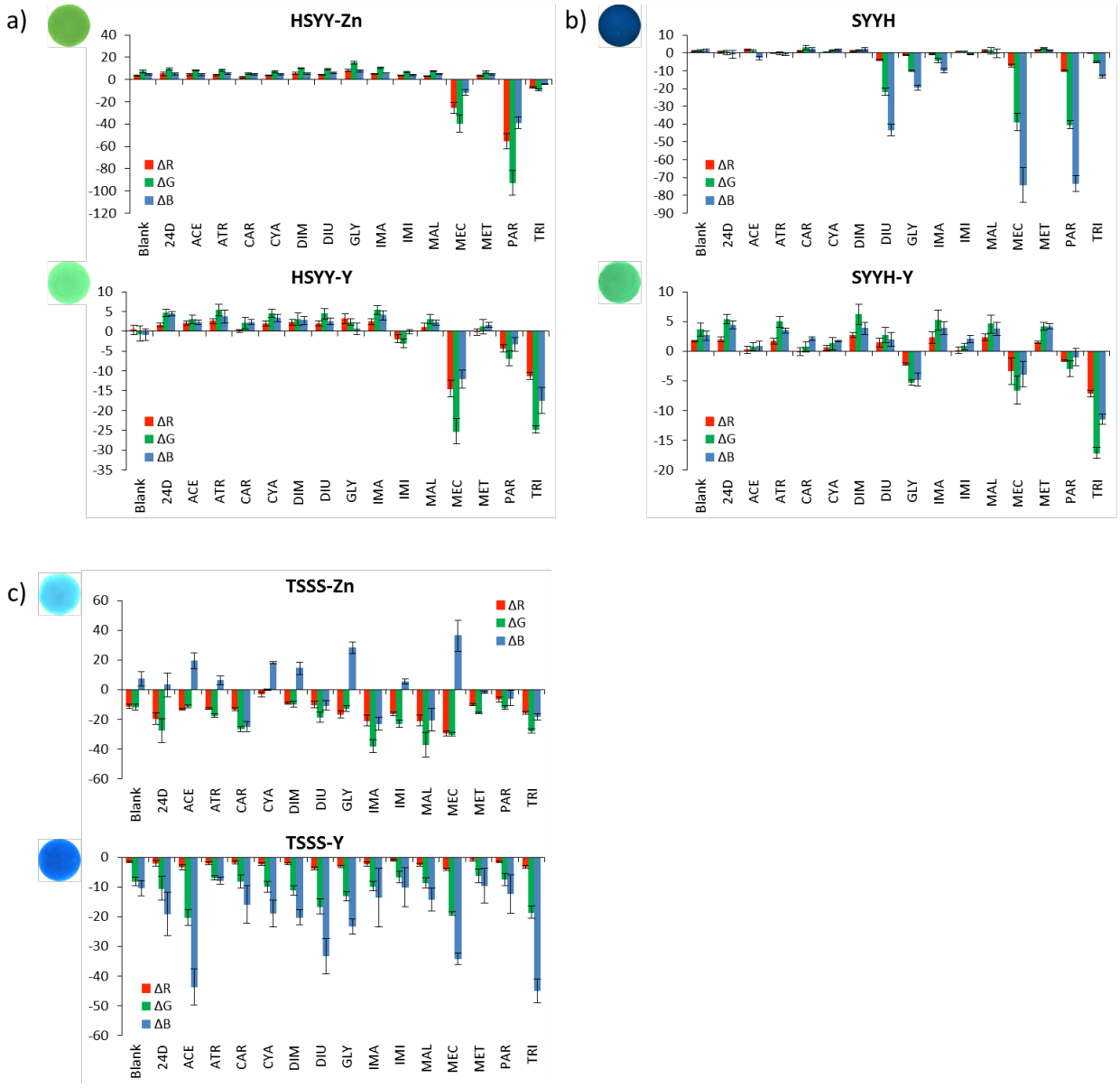


Figure S6. Comparison of responses of identical chemosensors with different metals. A sample bead image for each sensor sequence is also included. This data was collected from the data set represented in Figure S4.

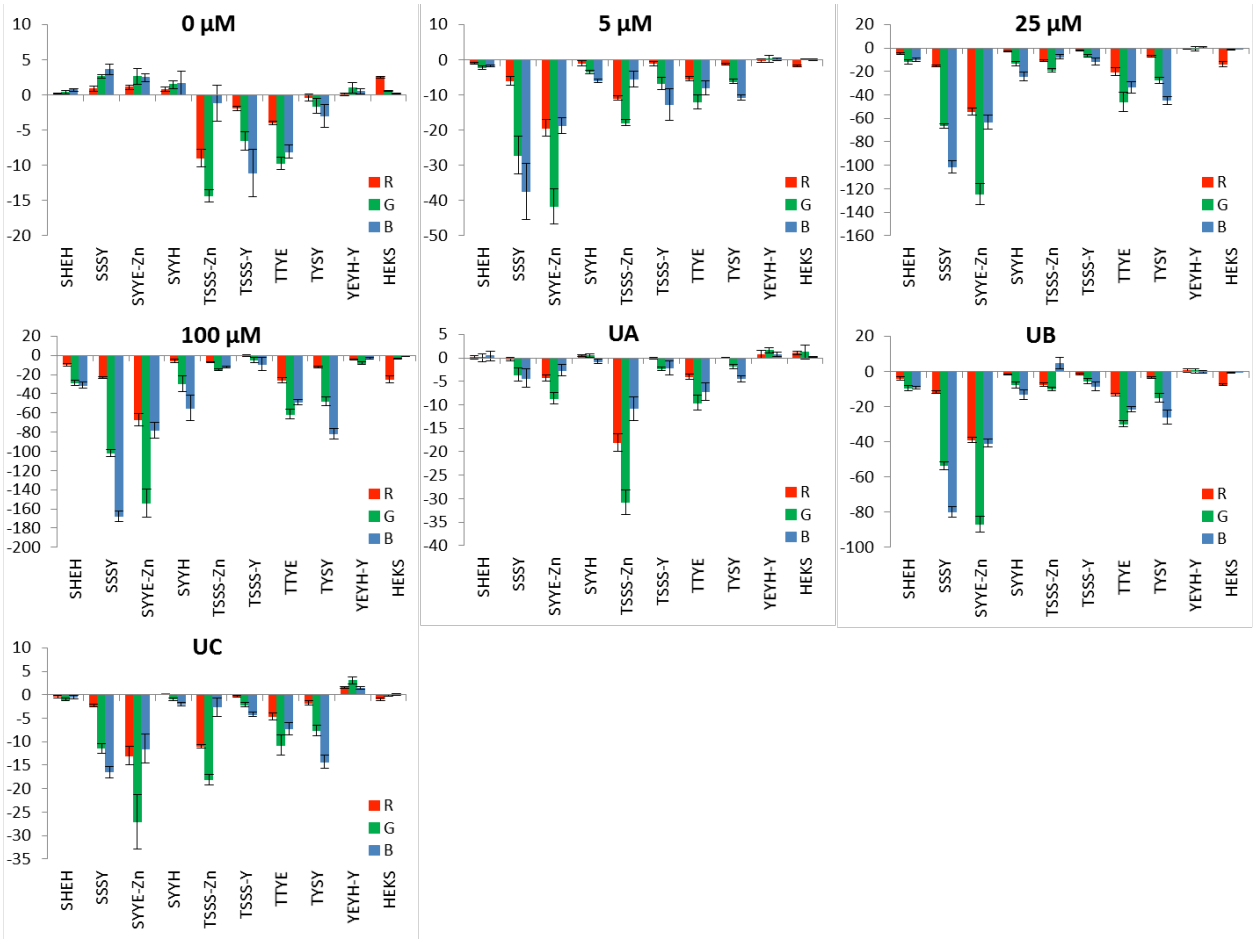


Figure S7. Data used for estimating concentrations of paraquat in unknown samples. Quantitative fluorescence responses (ΔRGB from microscopy images) of the ten-sensor set from unknown concentration tests of paraquat in buffered water (value falling in between standards; from 0 to 100 µM). Red, green, and blue bars represent changes in R, G, and B (on a scale of ±255), respectively, and the error bars represent standard deviation. The unknown concentrations were: UA = 1 µM, UB = 15 µM, and UC = 2 µM.

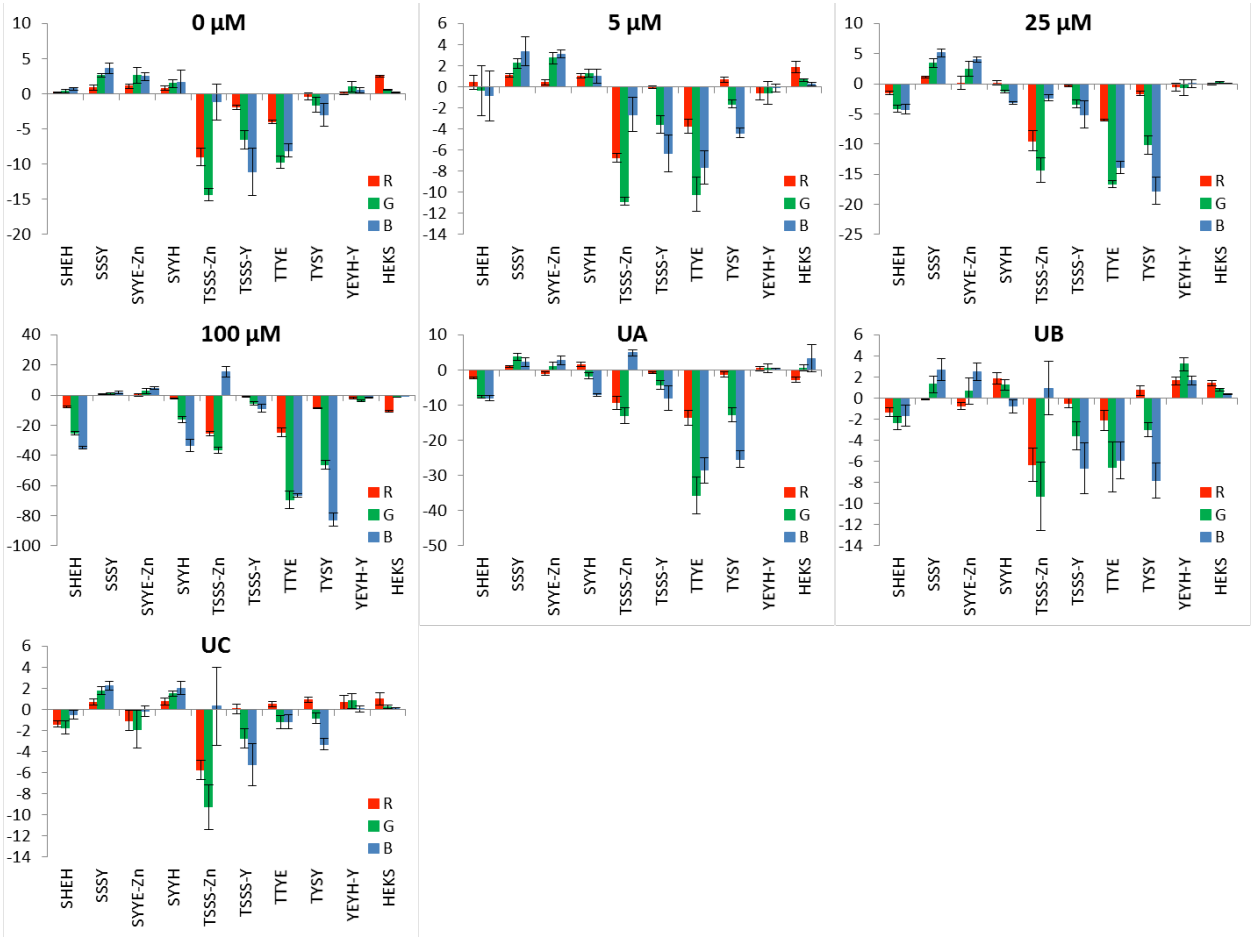


Figure S8. Data used for estimating glyphosate concentrations in unknown samples. Quantitative fluorescence responses (ΔRGB from microscopy images) of the ten-sensor set from unknown concentration tests of glyphosate in buffered water (value falling in between standards; from 0 to 100 µM). Red, green, and blue bars represent changes in R, G, and B (on a scale of ±255), respectively, and the error bars represent standard deviation. The unknown concentrations were: UA = 40 µM, UB = 10 µM, and UC = 3 µM.

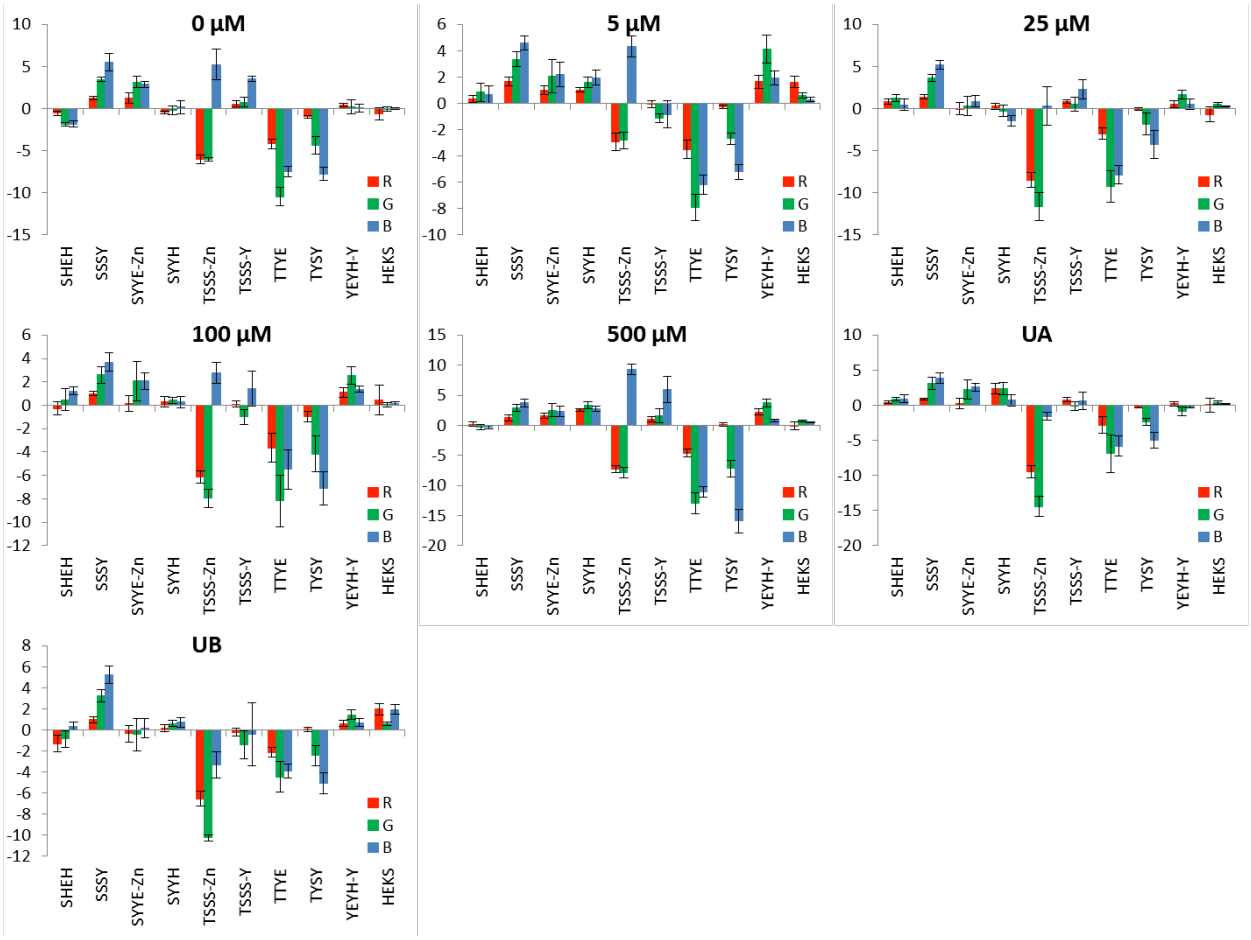


Figure S9. Data used for attempted estimation of dimethoate concentrations in unknowns. Quantitative fluorescence responses (ΔRGB from microscopy images) of the ten-sensor set from unknown concentration tests of dimethoate in buffered water (value falling in between standards; from 0 to 500 µM). Red, green, and blue bars represent changes in R, G, and B (on a scale of ±255), respectively, and the error bars represent standard deviation. The unknown concentrations were: UA = 50 µM and UB = 15 µM.

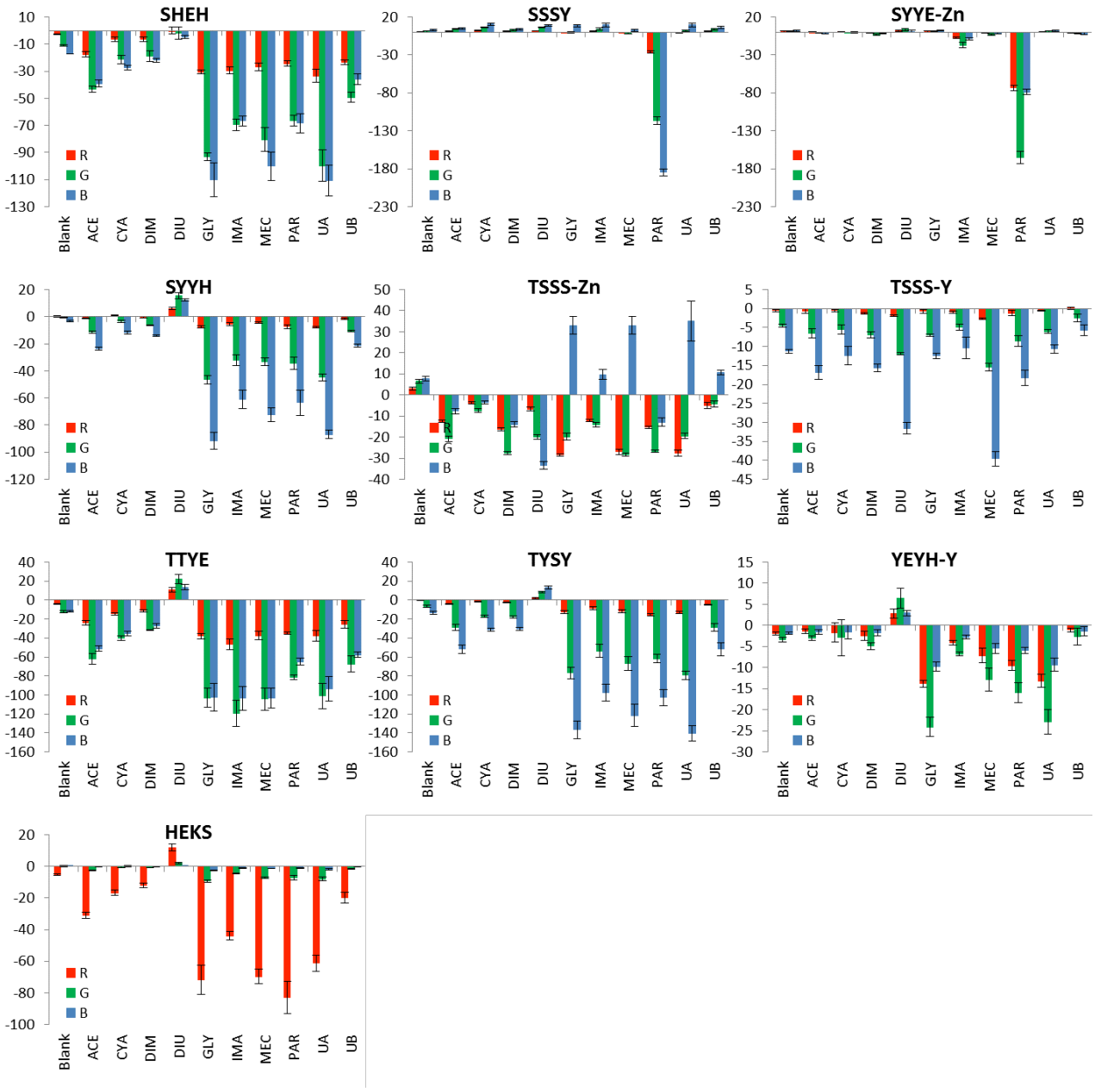


Figure S10. Blind identification of two unknown pesticides in lake water. Quantitative fluorescence responses (ΔRGB from microscopy images) of the ten-sensor set from two unknown pesticides in lake water, with comparison to eight known standards (100 µM) in the same water (Felt Lake, Portola Valley, California). Red, green, and blue bars represent changes in R, G, and B (on a scale of ±255), respectively, and the error bars represent standard deviation. The unknown pesticides were: UA = glyphosate (GLY) and UB = acephate (ACE).

References

1. L. H. Yuen, R. M. Franzini, S. Wang, P. Crisalli, V. J. Singh, W. Jiang, E. T. Kool, *Angew. Chem. Int. Ed.* **2014**, *53*, 5361-5365.
2. R. X. F. Ren, N. C. Chaudhuri, P. L. Paris, S. Rumney IV, E. T. Kool, *J. Am. Chem. Soc*. **1996**, *118*, 7671-7678; J. Gao, C. Strässler, D. Tahmassebi, E. T. Kool, *J. Am. Chem. Soc*. **2002**, *124*, 11590-11591.
3. S. Wang, J. Guo, T. Ono, E. T. Kool*, Angew. Chem. Int. Ed*. **2012**, *51*, 7176-7180.
4. H. Kwon, W. Jiang, E. T. Kool, *Chem. Sci*. **2015**, *6*, 2575-2583.
5. M. H. Ohlmeyer, M. Swanson, L. W. Dillard, J. C. Reader, G. Asouline, R. Kobayashi, M. Wigler, W. C. Still, *Proc. Natl. Acad. Sci. U. S. A*. **1993**, *90*, 10922-10926; H. P. Nestler, P. A. Barlett, W. C. Still, *J. Org. Chem*. **1994**, *59*, 4723-4724.