Published in final edited form as:

Chem Sci. 2013 August 1; 4(8): 3184–3190. doi:10.1039/C3SC50985K.

DNA-polyfluorophore Chemosensors for Environmental Remediation: Vapor-phase Identification of Petroleum Products in Contaminated Soil†

Wei Jiang, Shenliang Wang, Lik Hang Yuen, Hyukin Kwon, Toshikazu Ono, and Eric T. Kool

Department of Chemistry, Stanford University, Stanford, California, 94305-5080, United States.

Fax: +650 725 0259; Tel: +650 724 4741

Eric T. Kool: kool@stanford.edu

Abstract

Contamination of soil and groundwater by petroleum-based products is an extremely widespread and important environmental problem. Here we have tested a simple optical approach for detecting and identifying such industrial contaminants in soil samples, using a set of fluorescent DNA-based chemosensors in pattern-based sensing. We used a set of diverse industrial volatile chemicals to screen and identify a set of five short oligomeric DNA fluorophores on PEG-polystyrene microbeads that could differentiate the entire set after exposure to their vapors in air. We then tested this set of five fluorescent chemosensor compounds for their ability to respond with fluorescence changes when exposed to headgas over soil samples contaminated with one of ten different samples of crude oil, petroleum distillates, fuels, lubricants and additives. Statistical analysis of the quantitative fluorescence change data (as $\Delta(R,G,B)$ emission intensities) revealed that these five chemosensors on beads could differentiate all ten product mixtures at 1000 ppm in soil within 30 minutes. Tests of sensitivity with three of the contaminant mixtures showed that they could be detected and differentiated in amounts at least as low as one part per million in soil. The results establish that DNA-polyfluorophores may have practical utility in monitoring the extent and identity of environmental spills and leaks, while they occur and during their remediation.

Introduction

Contamination of soil and groundwater by petroleum-based products is an extremely important and widespread environmental problem. Such contamination arises from above-ground leaks in tanks and pipelines, from degradation of aging underground storage tanks, and from breaches that occur during production and shipping.^{1,2} Petroleum-based mixtures can be acutely toxic to animals, plants and humans,³ and long-term exposure may carry health risks as well.⁴ The magnitude of the problem is large; for example, in the U.S. there are an estimated 80,000 underground tanks currently leaking into the surrounding soil, with ca. 6,000 more discovered each year.⁵ The spreading “plumes” of organic contaminants put the associated groundwater at risk.⁶

As a result of this widespread hazard, remediation of such spills is of high priority worldwide.^{7,8} Analysis is important to determine their makeup and origin, and to monitor the margins of each spill as it is being cleaned or removed. Laboratory testing methods can

†Electronic supplementary information (ESI) available: Synthesis and characterization data; details of experimental methods; color-change plots and scattering data. DOI: 10.1039/b000000x

Correspondence to: Eric T. Kool, kool@stanford.edu.

be quite effective in determining chemical compositions of organic mixtures in soil samples,⁹ but the analysis typically requires expensive instrumentation such as gas chromatography-mass spectrometry and trained technicians and analytical chemists to operate them. This adds time and cost to testing soil samples, and limits the ability to quickly assess the progress of cleanup efforts.

As a result, there is a need for analytical methods that can quickly assess levels of contaminants in soils in the field, and differentiate their compositions. Among the most promising approaches for characterizing vapors from samples are optical methods, using compounds or materials that respond to vapors of analytes with changes in reflectance, absorbance or fluorescence.^{10–15} Optical methods are rapid and quantifiable, and may be implemented with very small amounts of sensor compounds. Arrays of compounds can be used in pattern-based responses to differentiate one analyte from another.^{10–15} However, although chromophore arrays have recently been used to analyze mixtures associated with foods^{16–18} and bacteria,^{19,20} we know of few, if any, tests of such optical methods in petroleum-based contamination of soil.²¹ This may be challenging for two reasons: first, such contamination may be diluted by large volumes of soil; and second, differentiating distinct petroleum-based products may be difficult because of the complexity of the mixtures, the similarity of many petroleum products, and the lack of chemical functionality of the main hydrocarbon components.

In previous studies we have investigated the use of DNA-polyfluorophores as fluorescent chemosensors.^{14,22,23} Our structural design involves the use of the DNA backbone as a scaffold to arrange multiple fluorophores in close proximity. The DNA phosphodiester backbone offers the advantages of rapid automated synthesis of thousands of possible optically distinct molecules from a set of monomer components, and serves to place the fluorophores in direct physical contact, as the nucleobases are in DNA. This proximity allows the fluorophores to yield sequence-dependent emergent optical properties, including several forms of energy/excitation transfer, that do not occur in the monomer components.²⁴ Such polyfluorophore chemosensors have been studied for their ability to respond with fluorescence changes to vapors of pure organic compounds and to toxic gases.^{22,23} Several sequences have also been used in combination to sense mixtures such as bacterial metabolites in air.²⁰ Despite these early examples, it remained to be seen whether such compounds could respond in a meaningful way to hydrocarbon-based complex mixtures, which contain less chemical functionality than previous mixtures. Moreover, it was not known whether they would have sufficient sensitivity to respond to such mixtures diluted to a large extent such as might occur in contaminated soil.

Here we describe the study of such DNA polyfluorophores (also termed oligodeoxyfluorosides or ODFs, Fig. 1) as possible optical chemosensors for detection and differentiation of closely related petroleum-based products in contaminated soil. The polyfluorophores were conjugated to PEG-polystyrene beads in a combinatorial library, and by screening a set of distinct petrochemical organic vapors we arrived at five ODF sequences that gave strong and highly diverse responses. We then studied this set of ODFs for their pattern-based responses with a set of ten petroleum-based products, including gasolines, crude oil, and oil-based additives. We used statistical methods to analyze the quantitative fluorescence responses, which ranged from quenching to lighting-up and wavelength shifts. Experiments showed that these five ODFs could successfully differentiate all ten products despite many similar hydrocarbon compositions, and they could detect and differentiate contaminants at concentrations as low as one milligram per kilogram of soil (1 ppm) within 30 minutes.

Experimental Methods

Phosphoramidite monomers and ODF library construction

The syntheses of the deoxyriboside monomers Y, E, V, Q, F, K were performed as previously reported.^{25–28} The spacer phosphoramidite (S) and the 5,6-dihydro-dT-CE phosphoramidite (H) were purchased from Glen Research. The construction of the tetramer library was carried out by standard split-and-pool methods on 130 mm amine-functionalized PEG-polystyrene beads (NovaSyn TG amino resin, Novabiochem) to yield 4096 ODF sequences. Individual beads were chemically tagged for sequencing by the method of Still.²⁹

ODF oligomers on beads

The 20 selected sensor sequences from initial screening were resynthesized on an ABI 394 DNA synthesizer with standard phosphoramidite chemistry. The synthesis was performed at 1 mmole scale in a standard column containing both PEG-PS beads and 3'-Phosphate CPG (Glen Research), so that sensor beads and corresponding ODFs on standard CPG were made at the same time. This allowed the characterization of the ODFs off of the CPG beads while making the corresponding sensors on PEG-PS beads in one synthesis. Cleavage of ODF oligomers from CPG beads were carried out with 0.05 M K₂CO₃ in methanol (room temperature, 8 hours). Then they were purified by reverse-phase HPLC and characterized by MALDI-MS and optical spectra (see Electronic Supplementary Information (ESI)).

Library screening

To screen potential sensors of VOCs, PEG-PS beads of the library were placed on a small microscope slide and enclosed in a sealed 5 mL quartz fluorescence cuvette (QS 111, Hellma). Before screening, images were taken in air under an epifluorescence microscope with 4X objective ($\lambda_{\text{ex}} = 340\text{--}380\text{ nm}$; $\lambda_{\text{em}} > 420\text{ nm}$) with a Spot RT digital camera and Spot Advanced Imaging software. After this, one drop of VOC liquid or 10 mg of VOC solid was placed beside the microscope slide in the cuvette, which was sealed to generate saturated vapors. After 30 min exposure at room temperature, fluorescence images were taken again. The images were analyzed with Adobe Photoshop by constructing 50% gray-based difference maps of the beads by merging the inverse of “before” images with “after” images using 50% transparency (Fig. S1). Beads that gave the strongest responses were picked up with a flame-pulled pipette and transferred into a capillary tube for decoding by release of the chemical tags, which were analyzed by electron-capture gas chromatography.²⁹

Petroleum-based analytes

Analytes were STP Gas Treatment, Turbo 108 Octane Booster, Prestone High Temperature Synthetic Brake Fluid, 3-IN-ONE Multi-purpose Oil, Clean-Strip Kerosene heater fuel (Home Depot), 87-octane gasoline (Valero), 91-octane gasoline (Shell), E85 gasoline (Propel) and diesel fuel (Valero). Crude oil was Colorado Sweet Crude (eBay).

Sample preparation and screening

For the preparation of soil samples contaminated with petroleum-based analytes, crude soil (Earthgro top soil, Home Depot, 40 lb bag) was baked in an oven (ca. 200 °C) for 2 days to remove excess water, then filtered with a sieve. The fine soil was then placed at room temperature open to air for 3 days before use. When doing a sensing experiment, a calculated amount of soil was weighed and placed in a 100 or 1000 mL round-bottom flask sealed with a rubber septum. After 30 minutes' equilibration in the flask, 5 mL of vapor above the soil was extracted by gas-tight syringe and injected in a sealed 5 mL quartz Hellma QS 111 cuvette containing sensor beads. The cuvette was kept at room temperature

for 30 minutes and then a fluorescence image was taken as blank background. Then a calculated amount of petroleum analyte was added into a second weighed soil sample in a similar round-bottom flask by pipette and the sealed bottle was shaken vigorously to evenly mix the sample. The sample bottle was equilibrated at room temperature for another 30 minutes. 5 mL of the contaminant vapor was extracted and injected in the sensor-containing cuvette which was pre-evacuated under high vacuum. Then the cuvette was vented with a syringe needle and another 10 mL of the contaminant vapor was continuously injected via syringe pump during 30 minutes to keep the concentration of vapor consistent. A fluorescence image of the beads was taken 30 minutes after exposure to the contaminant vapor. Care was taken not to disturb the beads between images.

Optical and statistical analysis

Fluorescence change data from bead images were analyzed with Adobe Photoshop by reading the RGBL (red, green, blue, luma) values of the subtracted images before and after analyte exposure. Average RGBL values were determined from a 16×16 pixel box in the center of each bead image. Quantitative color changes after exposure to VOCs, expressed in ΔR , ΔG , ΔB and ΔL on a ± 256 unit scale, were calculated for each bead and averaged (no less than three beads per analyte for a given sensor). Standard deviations and standard errors were determined to evaluate the accuracy and reproducibility of the responses. The statistical calculations, principal component analysis (PCA) and agglomerative hierarchical clustering (AHC) analysis, were performed with Addinsoft XLSTAT software using ΔR , ΔG and ΔB values as input.

Results

DNA-polyfluorophore design and composition

The DNA-polyfluorophores in this study were composed of various combinations of eight different monomer fluorophore nucleosides (**Y**, **E**, **V**, **F**, **Q**, **K**) and spacer groups (**S**, **H**) (Fig. 1A). The goal in choosing the monomers was to yield a range of colors across the spectrum, to enable changes in the blue, green, and red channels. For simplicity and small size, we limited lengths to tetramers; this yields 4096 unique sequences, which are denoted by letters describing the monomers in 5'→3' direction, analogous to DNA. The library was constructed on 130-micron PEG-polystyrene beads by standard split-and-mix methodology, and sequences were encoded on the beads with chemical tags. The library of sequences could be readily imaged by epifluorescence microscopy (Fig. 1C); for imaging we used a single excitation filter (340–380 nm) and observed all visible fluorescence in real time (RGB camera) with a 420 nm long-pass filter. The library showed a broad range of fluorescence properties in air, with emission colors ranging from blue to white to red and varying greatly in brightness.

Selection of a set of ODFs for differentiation of industrial vapors

Our initial goal was to identify a preliminary set of ODF sequences that could respond broadly and with good intensity to a wide range of possible industrial organic contaminants, including many petroleum components and petrochemicals. Thus we chose a set of twelve organic volatile compounds, including petroleum components (alkanes, alkenes, aromatics) and halogenated hydrocarbons (Table 1 and Table S2). These were screened one at a time with the ODF library by simply placing a drop next to beads in a chamber, yielding near-saturated vapors in air. Changes in fluorescence after 30 min were detected by before/after difference maps (see Fig. S1 in ESI); beads that gave the greatest changes were picked and sequenced. We arrived at a set of twenty ODF sequences that each responded strongly to at least one of the twelve compounds. These twenty sequences on beads were re-synthesized

individually and characterized by absorption and emission spectra and MALDI-mass spectrometry (Table S1 and Fig. S2).

We then set out to narrow this set of twenty candidate chemosensors to a small minimal set by carrying out cross-screening experiments, measuring quantitative responses of each ODF to each of the twelve petrochemical compounds. Graphical maps of the cross-screening data for the 240 combinations are given in Figs. S3 and S4, and quantitative R,G,B response data were recorded with statistical software. We used Principal Component Analysis (PCA) and Agglomerative Hierarchical Clustering (AHC) to group the candidate chemosensors for their diversity of response and ability to differentiate the analytes (Figs. 2,3 and S5). Using the clustering analysis we then chose varied sets of five sensor sequences from dissimilar classes (Fig. 3 and S12) and statistically compared their ability to differentiate the industrial analytes. This resulted in a final optimal set of five ODF sequences that best differentiated the VOC compounds: YSES, EYYY, YQFS, FEYF, and SSEK (Fig. S7; see three-dimensional scattering in Movie_S1). Other sets of five also successfully differentiated the twelve analytes (Fig. S12), but this set gave the largest scattering of responses overall.

Differentiating petroleum-based mixtures in contaminated soil

The above experiments determined a set of five chemosensor ODFs that responded with diverse optical signatures to individual components of oil and to petroleum-derived compounds. We then used this set of chemosensors to analyze petroleum-based products in soil, representative of the varied types of mixtures that might be found in samples taken during remediation of spills and leaks. To this end, we developed a simple standardized approach, mixing measured amounts of a given product in soil and allowing it to equilibrate in a glass flask with air above the sample. This headgas was then injected by slow flow (30 min) through a quartz optical cell containing the ODF beads (see Experimental section and Fig. S6 in ESI). Once again, digital images of beads before (uncontaminated soil only) and after exposure to contaminated soil were subtracted, and changes were quantified as ΔR , ΔG , ΔB data.

The petroleum-based mixtures tested included 87- and 91-octane gasoline, E85 gasoline, diesel fuel, kerosene, crude oil, multipurpose oil, brake fluid and two commercial gasoline additives (Table 1). Experiments were initially carried out at 1000 ppm in soil (i.e., 10 mg in 10 g soil), and all ten mixtures were individually tested with the five ODF sequences, with three replicates each to allow for a measure of error. The ΔR , ΔG , ΔB data are given in Figs. S7,S8, and statistical plots are shown in Fig. S9. Examination of the R,G,B changes showed that all the mixtures gave similar overall responses, as expected since nearly all of the mixtures are closely related in main composition. Responses varied widely with ODF sequence; for example, YQFS yielded light-up responses for most analytes, while EYYY commonly gave quenching responses (Fig. S8). Overall, error bars were small, showing good precision and repeatability, and this allowed small differences among the analytes to be evaluated quantitatively. The PCA analysis of the multidimensional data showed good scattering of all ten analyte mixtures (Fig. S9). Notably, the first two principal component axes (F1 and F2) accounted for only 67% of the scattering data, showing high dimensionality of the responses of the five ODFs. Thus data were better represented in 3D than in 2D plots; a 3D video is available (Movie_S2). A dendrogram of the responses generated by AHC analysis is shown in Fig. S9. The most similar responses were seen for crude oil and brake fluid; next most similar were diesel fuel and one gasoline additive. The most dissimilar responses (i.e., the most widely resolved) were seen for the crude oil/multipurpose oil samples relative to the refined gasolines. Importantly, measurements showed high repeatability and precision: all repeats of measurements (taken with sensors on different beads) grouped together with very low dissimilarity.

Testing limits of detection in soil

Finally, we carried out measurements to evaluate the ability of the five chemosensors to detect such mixtures at lower concentrations in soil. To carry this out we selected three of the petroleum products and measured responses of the five-sensor set at further tenfold dilutions in soil: 10 ppm and finally 1 ppm. The three analytes tested in this way were kerosene, E85 gasoline, and 91-octane gasoline. These were again measured with three replicates each to obtain a measure of the precision and repeatability. The RGB response data for the 1000, 100, and 10 ppm contaminants are given in Fig. S10; comparison shows, as expected, that the higher the concentration, the larger the responses are. However, the 10 ppm analytes were all readily detected with nonzero responses and were separated from each other (ESI). We then proceeded to analyze the 1 ppm data, which are plotted in Fig. 4. The results show that the first three dimensions of the principal component analysis yield clear separation of the three analytes. Clustering analysis (Fig. 5) shows that all three are easily resolved from uncontaminated soil. Thus the data show that the five ODFs are able to detect and identify three related petroleum contaminants in soil at levels at least as low as one part per million.

Discussion

Our data show that a minimal set of five DNA-like ODF fluorophores on beads was able to differentiate all ten of the petroleum-based contaminants in soil. Examination of the dendrograms of the AHC analysis from the fluorescence responses (Fig. S9) shows the relationships of similarity/dissimilarity after the measurements. The data show that repetitions of the fluorescence measurements in all cases grouped a given product with another replicate of the same product (see also Movie_S2), giving high confidence that the distinct mixtures were well differentiated from one another in a repeatable way. Indeed, the level of discriminating ability is surprising, considering the closely related compositions of some of the analytes tested here. For example, kerosene, 91-octane gasoline, 87-octane gasoline, and diesel fuel are very similar in composition,³⁰ yet these four mixtures were differentiated with high confidence.

Although the compositions of all the commercial products we tested are not publicly known, our statistical groupings allow us to categorize them by similarity to other mixtures. One example of such an unknown sample is the gas treatment fluid, which showed closest similarity to diesel fuel and 87-octane gasoline, suggesting that its base composition may be close to these. It is not surprising that the less-refined oil samples and E85 fuel were strongly differentiated from the refined gasoline-like products, as the mixtures are no doubt chemically quite distinct. However, it was unexpected how well-differentiated the 91-octane gasoline was from the 87-octane gasoline and diesel fuel. It seems possible in this case that the chemosensors may be responding to an additive in the high-octane blend that is not present in the others. Overall, the results highlight the fact that pattern-based sensing is especially adept at comparing complex analytes, even when their compositions are not known.

The current set of ODF chemosensors was chosen based on its ability to differentiate a wide range of single chemical components that can be found in petroleum and other industrial spills. This approach is expected to yield a set of dyes that can respond with versatility to a wide range of different industrial spills, although more testing would be needed with non-petroleum mixtures to establish whether they would be as effective as they were here. It seems possible that greater separating ability might be possible for specific analyte mixtures if desired (for example, stronger differentiation of diesel from high-octane gasoline); this might be achieved by screening a library of ODFs directly with specific analyte mixtures rather than single compounds. Future experiments could readily test this hypothesis.

The ODF fluorophores studied here were highly sensitive to the current petroleum contaminants. Our data showed that three of the contaminant mixtures were readily detected at 1 ppm in soil (the equivalent of one milligram per kilogram). For reference, New York State standards for gasoline-contaminated soil places “Alternative Guidance Value” limits on most individual components (e.g., MTBE, xylene, naphthalene) at 0.1–1 ppm.³¹ Since we detected total gasoline mixtures at 1 ppm, the components were at much lower concentrations than these guidance values. We attribute the high sensitivity to multiple possible factors: first, the use of fluorescence allows for inherently high sensitivity and wide dynamic range of response. Second, the placement of the ODFs on PEG-polystyrene beads also likely enhances response by concentrating the nonpolar volatiles near the fluorophores; consistent with this notion, we did observe evidence of bead swelling in the most concentrated samples tested here. Third, the large aromatic surfaces in the ODFs may interact favorably with many aromatic components in petroleum. It would certainly be of interest in the future to test responses of these sensor compounds to petroleum contaminants at levels yet lower than 1 ppm in soil. However, in the practical sense it becomes difficult (in laboratory experiments) to prepare samples of lower concentration than this, as the liquids become difficult to measure and dispense at volumes below 500 nanoliters, and soil samples in the laboratory become increasingly unwieldy to handle in amounts above 500 g. Nevertheless, with more sophisticated mixing and dilution strategies (and higher-volume equipment) such experiments may be of interest in the future, as workers involved in remediation may wish to measure contaminant concentrations lower than 1 ppm at the margins of a spill.

The pattern-based approach to optical sensing offers a number of practical advantages over other approaches for analysis of complex mixtures. The use of fluorescence changes gives a large degree of data (here, a ± 256 -unit scale at three different wavelength ranges); this complexity in five different sensor compounds gives a large potential range of multidimensional responses (ca. 10^{40}). As a result, one can analyze a large number of distinct samples with a small number of sensor molecule structures. Second, as mentioned above, it is important to note that in this pattern-based approach there is no need to know *a priori* what the chemical composition is of a given analyte mixture. Indeed, one may never require this information, since differences can be distinguished reproducibly by the distinct optical responses. On the other hand, if one does desire to know the chemical composition, one can identify specific mixtures simply by testing prior training samples to define the expected pattern in the chemosensors. A third benefit of using optical pattern-based responses is that they can be extremely effective at differentiating closely related mixtures. Recent examples of this include differentiating coffee¹⁷ and wine samples,¹⁶ identifying different classes of cultured bacteria,^{19,20} and differentiating petroleum mixtures in the current study. A final advantage of this approach is its simple practical implementation, using disposable chemosensors and measuring optical changes in a simple way that could be readily automated and taken to the field. This may incur lower cost and require potentially less training than current methods; moreover, it also allows for portability, which is significant because the user could analyze fresher samples containing more volatiles that might otherwise be lost during storage and transport.

More work will be helpful in addressing practical issues in optical sensing before ODFs can be employed in remediation of authentically contaminated soil samples. For example, development of methods for arraying the chemosensors on fixed surfaces would aid in array/pattern standardization relative to the current use of beads randomly scattered on slides. Second, it will be of interest to assess specifically how the headgas over a soil sample is optimally sampled and flowed over the sensors, since this may well affect sensitivity. Future work will be directed at these issues.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge Eni S.p.A. and the U.S. National Institutes of Health (GM067201) for support. We also thank S. K. Edwards and E. M. Harcourt for assistance with HPLC and fluorescence measurements.

Notes and references

1. Stalcup D, Yoshioka G, Mantus E, Kaiman B. *Int Oil Spill Conf Proc.* 1997;939–940.
2. Etkin DS. *Int Oil Spill Conf Proc.* 2001;1291–1300.
3. (a) Reese E, Kimbrough RD. *Environ Health Perspect.* 1993; 101:115–131. [PubMed: 8020435] (b) Seymour FK, Henry JA. *Hum Exp Toxicol.* 2001; 20:551–562. [PubMed: 11926609]
4. Lynge E, Andersen A, Nilsson R, Barlow L, Pukkala E, Nordlinder R, Boffetta P, Grandjean P, Heikkilä P, Hörte LG, Jakobsson R, Lundberg I, Moen B, Partanen T, Riise T. *Am J Epidemiol.* 1997; 145:449–458. [PubMed: 9048519]
5. Environmental Protection Agency. Annual Report on the Underground Storage Tank Program. 2011. http://www.epa.gov/oust/pubs/fy11_annual_ust_report_3-12.pdf
6. Shih T, Rong Y, Harmon T, Suffet M. *Environ Sci Technol.* 2004; 38:42–48. [PubMed: 14740715]
7. Douglas GS, McCarthy KJ, Dahlen DT, Seavey JA, Steinhauer WG, Prince RC, Elmendorf DL. *J Soil Contam.* 1992; 1:197–216.
8. Fountain, H. *New York Times.* 2010 Jun 24. Advances in Oil Spill Cleanup Lag Since Valdez.
9. Cam D, Gagni S. *J Chromatog Sci.* 2001; 39:481–486. [PubMed: 11718309]
10. Freund MS, Lewis NS. *Proc Natl Acad Sci U S A.* 1995; 92:2652–2656. [PubMed: 11607521]
11. Dickinson TA, White J, Kauer JS, Walt DR. *Nature.* 1996; 382:697–700. [PubMed: 8751439]
12. Janzen MC, Ponder JB, Bailey DP, Ingison CK, Suslick KS. *Anal Chem.* 2006; 78:3591–3600. [PubMed: 16737212]
13. Umali AP, Anslyn EV. *Curr Opin Chem Biol.* 2010; 14:685–692. [PubMed: 20801075]
14. Samain F, Ghosh S, Teo YN, Kool ET. *Angew Chem Int Ed.* 2010; 49:7025–7029.
15. Le ND, Rana S, Rotello VM. *Expert Rev Mol Diagn.* 2013; 13:111–113. [PubMed: 23477550]
16. Umali AP, LeBoeuf SE, Newberry RW, Kim S, Tran L, Rome WA, Tian T, Taing D, Hong J, Kwan M, Heymann H, Anslyn EV. *Chem Sci.* 2011; 2:439–445.
17. Suslick BA, Feng L, Suslick KS. *Anal Chem.* 2010; 82:2067–2073. [PubMed: 20143838]
18. Kwon H, Samain F, Kool ET. *Chem Sci.* 2012; 3:2542–2549.
19. Carey JR, Suslick KS, Hulkower KI, Imlay JA, Imlay KR, Ingison CK, Ponder JB, Sen A, Wittrig AE. *J Am Chem Soc.* 2011; 133:7571–7576. [PubMed: 21524080]
20. Koo CK, Wang S, Gaur RL, Samain F, Banaiee N, Kool ET. *Chem Commun.* 2011; 47:11345–11347.
21. Chakraborty S, Weindorf DC, Zhu Y, Li B, Morgan CL, Ge Y, Galbraith J. *Environ Monit.* 2012; 14:2886–2892.
22. Samain F, Dai N, Kool ET. *Chem Eur J.* 2011; 17:174–183. [PubMed: 21207614]
23. Koo CK, Dai N, Samain F, Kool ET. *Chem Sci.* 2011; 2:1910–1917.
24. Teo YN, Wilson JN, Kool ET. *J Am Chem Soc.* 2009; 131:3923–3933. [PubMed: 19254023]
25. Gao J, Watanabe S, Kool ET. *J Am Chem Soc.* 2004; 126:12748–12749. [PubMed: 15469249]
26. Wang S, Guo J, Ono T, Kool ET. *Angew Chem Int Ed.* 2012; 51:7176–7180.
27. Ono T, Edwards SK, Wang S, Jiang W, Kool ET. *Nucleic Acids Res.* 2013 May 3. [Epub ahead of print].
28. Singh V, Wang S, Kool ET. 2013; 135:6184–6191.
29. Ohlmeyer MH, Swanson RN, Dillard LW, Reader JC, Asouline G, Kobayashi R, Wigler M, Still WC. *Proc Natl Acad Sci U S A.* 1993; 90:10922–10926. [PubMed: 7504286]

30. (a) Coleman WE, Munch JW, Streicher RP, Ringhand HP, Kopfler FC. *Env Contam Toxicol.* 1984; 13:171–178.(b) Thomas DH, Delfino JJ. *Groundwater Monitoring Remed.* Fall;1991 :90–100.
31. <http://www.dec.ny.gov/regulations/30902.html#Table1>

Figure 1. Chemosensor compounds in this study. (A) Monomer components of oligodeoxyfluorosides (ODFs), consisting of fluorophore deoxyribosides and nonfluorescent spacers. (B) Structure of a representative ODF chemosensor (sequence YQFS) conjugated to a PEG-polystyrene bead. (C) Image of ODFs on beads, taken by epifluorescence microscopy.

Figure 2. Scattering data for 12 pure VOCs by Principal Component Analysis (PCA). (A) Plot of data projected to the first two principal axes (F1 vs. F2). (B) Orthogonal plot (F1 vs. F3). See Movie_S1 for animated 3D representation of scattering along the three axes.

Figure 3. Dendrograms of Agglomerative Hierarchical Clustering (AHC) analysis of 20 chemosensor responses to the 12 VOC analytes. (A) Grouping of analytes by dissimilarity of responses; (B) Grouping of sensors by dissimilarity of their responses to all twelve analytes.

Figure 4. Scattering of data for sensing petroleum contaminants at 1 ppm in soil by Principal Component Analysis (PCA). 2-D projection of data (F1 vs. F3) is shown. Data were obtained from five chemosensors (YSES, EYYY, YQFS, FEYF, SSEK).

Table 1

Analytes in this study.

Pure VOC analytes used in screening	Petrochemical mixtures tested in soil
benzene	kerosene
xylene	gasoline (87 octane)
naphthalene	gasoline (91 octane)
phenanthrene	gasoline (E85)
styrene	crude oil (Colorado sweet)
thiophene	diesel fuel
n-decane	multi-purpose oil
1,2-dibromoethane	brake fluid
1,2-dichlorobenzene	gasoline additive
methyl t-butyl ether	gasoline octane booster
acrylamide	
acrylonitrile	