

CrossMark  
click for updatesCite this: *Chem. Sci.*, 2016, 7, 199

## An optoelectronic nose for identification of explosives†

Jon R. Askim, Zheng Li, Maria K. LaGasse, Jaqueline M. Rankin and Kenneth S. Suslick\*

Compact and portable methods for identification of explosives are increasingly needed for both civilian and military applications. A portable optoelectronic nose for the gas-phase identification of explosive materials is described that uses a highly cross-reactive colorimetric sensor array and a handheld scanner. The array probes a wide range of chemical reactivities using 40 chemically responsive colorimetric indicators, including pH sensors, metal–dye salts, redox-sensitive chromogenic compounds, solvatochromic dyes, and other chromogenic indicators. Sixteen separate analytes including common explosives, homemade explosives, and characteristic explosive components were differentiated into fourteen separate classes with a classification error rate of <1%. Portable colorimetric array sensing could represent an important, complementary part of the toolbox used in practical applications of explosives detection and identification.

Received 20th July 2015

Accepted 22nd September 2015

DOI: 10.1039/c5sc02632f

www.rsc.org/chemicalscience

### Introduction

Intensive research efforts have been made for the detection and identification of explosive compounds. Sensitivity, selectivity, speed, analyte scope, environmental tolerance, device size, and cost all factor heavily into the balance between ideal and practical analysis. Many methods for screening have been investigated, including single-target colorimetric tests,<sup>1</sup> ion-mobility spectrometry (IMS),<sup>1,2</sup> electronic noses,<sup>3,4</sup> and fluorimetry.<sup>5,6</sup> Despite the breadth of analytical methods available, portable methods still have significant room for improvement: single-target analyses are cumbersome when screening a wide range of analytes; IMS is relatively costly, often requires thermal or ionizing desorption of analytes, and is most useful primarily for nitro-organic detection; and traditional electronic nose technology suffers from sensor drift, poor selectivity and environmental sensitivity (*e.g.*, to changes in humidity or to interferences).<sup>3,7,8</sup> In comparison, colorimetric sensor arrays have a broad analyte response, good environmental tolerance, and high selectivity; they are also small, fast, disposable, and can be analyzed using inexpensive equipment.<sup>9–11</sup>

Colorimetric sensor arrays combine multiple cross-reactive colorimetric sensors that probe a wide range of analyte chemical properties,<sup>12–16</sup> including Lewis and Brønsted acidity/basicity, molecular polarity, and redox properties. Using a combination of broadly reactive and specifically targeted sensors, colorimetric sensor arrays have been successfully used to differentiate even among similar analytes within diverse

families, including toxic industrial chemicals,<sup>12,17</sup> oxidants,<sup>18</sup> complex mixtures,<sup>19–22</sup> and pathogenic bacteria and fungi.<sup>23–25</sup>

We report here the development of a new colorimetric sensor array and handheld reader for the identification of explosives and their components. Several new classes of colorimetric sensors were developed including cross-reactive metal–dye salts and other dyes designed to take advantage of the reactivity of carbonyl and nitro compounds. The resulting printed array had forty sensor elements mounted in a snap-together, disposable cartridge (Fig. 1).

Combined with the colorimetric sensor array, a hand-held reader permits rapid acquisition of low-noise colorimetric data (Fig. 1c). The hand-held reader makes use of a contact image sensor (CIS), a technology commonly used in business card scanners. The careful control of lighting, lack of moving parts, and insensitivity to vibration provides the reader improved signal to noise and faster scan rates compared to other digital imaging techniques;<sup>26</sup> signal-to-noise ratios in real-time chemical analysis are a factor of 3–10 higher<sup>26</sup> than currently used



Fig. 1 The optoelectronic nose. (A) The linear array of colorimetric sensors and disposable cartridge. Cartridge side view (7.9 × 2.8 × 1.0 cm). (B) Cartridge front view. (C) Handheld reader/analyzer (12.8 × 9.5 × 4.0 cm) based on a color contact line imager.

Department of Chemistry, University of Illinois at Urbana-Champaign, 600 S. Mathews Ave., Urbana, IL 61801, USA. E-mail: ksuslick@illinois.edu

† Electronic supplementary information (ESI) available: Sampling details, handheld reader details, additional array response data, PCA component score plots, <sup>1</sup>H-NMR of DMDNB and PETN. See DOI: 10.1039/c5sc02632f





Table 1 Array composition (top) and color-coded legend of sensor categories (bottom)<sup>a</sup>

Spot	Name	Spot	Name
1	FeCl <sub>2</sub> + Nile Red + TsOH	21	HgCl <sub>2</sub> + Bromophenol Blue
2	a-Naphthyl Red + TsOH	22	LiNO <sub>3</sub> + Bromocresol Green
3	Tetraiodophenolsulfonephthalein + TsOH	23	AgNO <sub>3</sub> + Bromocresol Green
4	Pyrocatechol Violet + TsOH	24	Pb(OAc) <sub>2</sub> + Disperse Red
5	Bromocresol Green + TsOH	25	Bromophenol Blue + TBAH
6	Methyl Red + TsOH	26	Methyl Red + TBAH
7	Bromocresol Purple + TsOH	27	Chlorophenol Red + TBAH
8	DNPH + Pararosanine + TsOH (conc. A)	28	Nitrazine Yellow + TBAH
9	DNPH + Pararosanine + TsOH (conc. B)	29	Bromothymol Blue + TBAH
10	DNPH + Pararosanine + TsOH (conc. C)	30	Thymol Blue + TBAH
11	FeCl <sub>2</sub> + Nile Red	31	m-Cresol Purple + TBAH
12	ZnTPP	32	N,N'-diphenyl-1,4-diphenyldiamine + TBAH
13	ZnTMP	33	tolidine + TBAH
14	CoTMP	34	o-dianisidine + TBAH
15	CdTPP	35	Nile Red + Matrix A
16	Bromophenol Red	36	Nile Red + Matrix B
17	Bromophenol Blue	37	Merocyanine 540
18	Nile Red	38	1-ethyl-4-(2-hydroxystyryl)pyridinium iodide
19	Acridine Orange Base	39	TBAH + Tropaeolin O (dye acts only as marker)
20	Zn(NO <sub>3</sub> ) <sub>2</sub> + Bromophenol Blue	40	Methylene Blue

What	How	Why
Redox-sensitive dyes	Fenton chemistry / other oxidation	Redox compounds
pH indicators	pH and Lewis acidity/basicity	Acidic/Basic compounds
DNPH spot	Brady's test	Ketones and Aldehydes
Porphyrins	Ligation, Lewis acidity/basicity	Ligands
Metal salts	Metal complexation	Ligands
Solvatochromic dyes	Solvatochromism	Solvents
Strong base	Meisenheimer adduct formation	Nitroaromatics
Locator spot	None intended	Fiducial marker (array edge detection)

<sup>a</sup> TsOH = *p*-toluenesulfonic acid (1 M in 2-methoxyethanol); TBAH = tetrabutylammonium hydroxide (40% in H<sub>2</sub>O); DNPH = 2,4-dinitrophenylhydrazine.

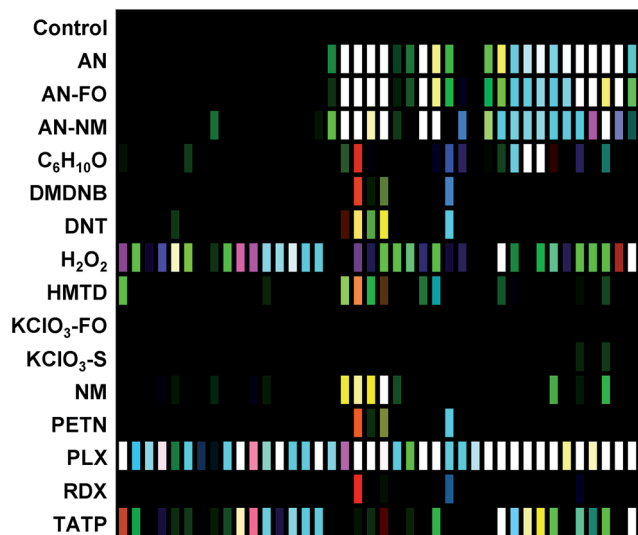


Fig. 3 Difference maps of the 40-element colorimetric sensor array showing signal-to-noise of 16 explosives, related analytes, and the control. S/N ratios of 3–10 were scaled for display on an 8 bit RGB color scale (i.e., 0–255).

(e.g., ammonium nitrate, AN); the array is detecting the volatile components of the analytes that form the “chemical bouquet”, which consists of volatile impurities and degradation products (in the case of AN, for example, these are generally ammonia and amines from the manufacturing process).

Limits of detection (LODs) were determined for the field-appropriate sampling protocol used in these studies and described in the Experimental section below. Air is pulled from a glass vial containing a mg-scale sample into the handheld reader using its internal micropump for 2 min as images of the sensor array were acquired. LODs for this sampling protocol were calculated using the single red, green, or blue response with the highest S/N for AN, NM, and DNT samples (representing highly responsive, moderately responsive, and weakly responsive analytes, respectively) using sample masses ranging from 0.5–100 mg; estimated LODs were as follows: AN = 0.32 mg, NM = 1.31 mg, DNT = 5.19 mg; cf. ESI p. S7.† While these mg-scale detection limits are by no means competitive with dedicated trace detectors, they are more than adequate for portable identification of HMEs under field conditions.



Improvement in sensitivity by the use of pre-concentrators is of course an option.

### Database evaluation

**Principal component analysis.** Principal component analysis (PCA)<sup>36–39</sup> was employed to provide an estimation of the dimensionality of the data acquired using the colorimetric sensor array, which is itself a measure of the dimensionality of the chemical reactivity space probed by the sensor array. PCA is an unsupervised statistical approach that generates a set of orthogonal vectors (*i.e.*, principal components) using a linear combination of array response vectors to maximize the amount of variance into the fewest number of principal components. The resulting principal components are *not* optimized for analyte discrimination: PCA describes the entire dataset using components to capture the maximum amount of total variance; this does not necessarily maximize discrimination ability among classes of analytes.

A scree plot of the normalized data collected for explosives analytes is given in Fig. 4. A total of 10 dimensions were required to capture 90% of the total variance and 16 dimensions for 95%; such high dimensionality is consistent with the very broad range of chemical reactivities being probed by the colorimetric sensor array, as we have noted before with other analytes.<sup>9,19,23,31</sup> The high dimensionality of the colorimetric sensor array is in stark contrast to traditional electronic nose technology in which only 1 or 2 dimensions are required to reach 95% of the total variance (in these cases, the sensor array is probing only a very limited range of chemical properties, *e.g.*, hydrophobicity/surface adsorption).<sup>9</sup>

When using PCA to discriminate among analytes, one typically plots the data points using the first two (or, rarely, three) principal components in a score plot. The assumption in these cases is that the vast majority of the discrimination ability is contained in these first few principal components. This is only true, generally, when the first few principal components also describe the vast majority of the total variance. The high dimensionality of the colorimetric sensor array data, however, makes PCA generally ill-suited for use in discrimination among

analyte species: too little of the variance is captured in two or three dimensions to avoid overlap of analyte classes, and in fact, PCA score plots show significant overlap among classes even among samples that show obvious qualitative differences in response (ESI, Fig. S6†).

**Hierarchical cluster analysis.** Hierarchical cluster analysis (HCA)<sup>36,38</sup> was used to give a model-free evaluation of the acquired database and the relative similarities among the data collected. HCA is an unsupervised clustering technique that groups array response data in the full 120-dimensional vector space (*i.e.*, color difference changes in red, green, and blue for each of the 40 sensor elements on the colorimetric sensor array); starting with single data points, clusters are formed hierarchically by connecting the centroids of unconnected clusters or data points (in this case using Ward's method, which minimizes the total within-cluster variance). The resulting dendrogram shows connectivity (indicating which clusters are most similar to each other), and inter-cluster distance (describing the magnitude of dissimilarity between clusters). The HCA dendrogram for the response of these common explosives is shown as Fig. 5. The method shows obvious discrimination among 11 of the 16 analytes (including the control). Confusions of clustering were observed among two groups: that containing the weakly-responding potassium chlorate mixtures (KClO<sub>3</sub>–fuel oil and KClO<sub>3</sub>–sugar) and

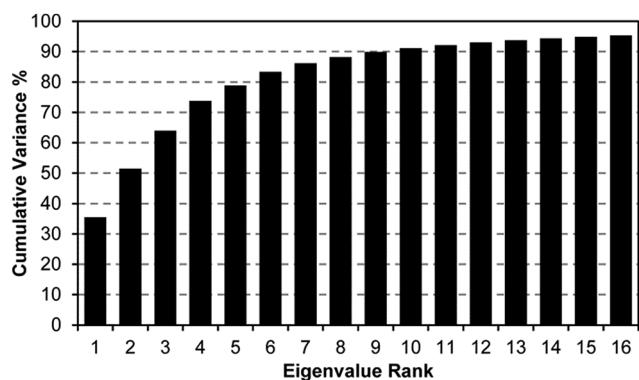


Fig. 4 Scree plot of the principal component analysis for 15 explosives and related compounds. 16 dimensions were required to capture 95% of the total variance, consistent with the colorimetric sensor array probing a wide range of chemical reactivity.

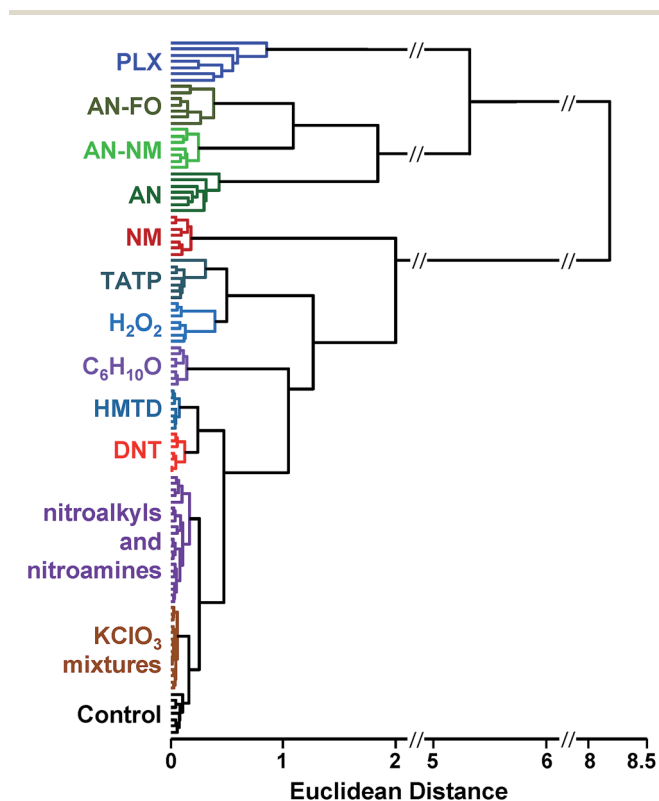


Fig. 5 Hierarchical cluster analysis (HCA) dendrogram of the normalized difference vectors (*i.e.*, changes in reflectance) for 16 explosives, related analytes, and the control; 112 trials in total. All species were clearly differentiable except among members of two groups: KClO<sub>3</sub> mixtures (KClO<sub>3</sub>–sugar and KClO<sub>3</sub>–fuel oil) and nitroalkyls/nitroamines (DMDNB, PETN, and RDX).





separately, the group containing nitroalkyls and nitroamines (PETN, RDX, and DMDNB).

**Customized HCA or PCA algorithms.** If one were insistent upon using HCA or PCA for discrimination, it is possible to use additional information (such as principal components or cluster centroids) to construct a high-dimensional space, and to develop a supervised algorithm to determine the minimum amount of information (*e.g.*, the number of principal components) required to discriminate with some arbitrary level of accuracy. Doing this, however, defeats the purpose of using HCA or PCA: these methods are fast, simple analytical tools that work well for initial determination of the quality of a dataset. There are better choices for determining discrimination ability in high-dimensional space, such as support vector machines, as described below.

### Classification methods

Classification methods involve developing *classifiers* – algorithms that can predict the identity of an unknown compared to an established database (*i.e.*, library or training set). Classifiers are based on a decision boundary by which an incoming sample can be classified, for example “Analyte A vs. Analyte B” or “Analyte C vs. not Analyte C”. Of necessity, such classifications are supervised methods: the identity of the samples in the database must be known. While unsupervised methods cannot be used to classify data, they can be used as the first step to develop the decision boundaries.

HCA and PCA are unsupervised methods used to analyze an existing data set. Because they are unsupervised, they provide no direct method for class prediction from data on an unknown sample (*i.e.*, classification). Although HCA and PCA do not present a direct method for classification, they can be used to develop decision boundaries if clustering is completely free of errors/confusions. As described above, however, these decision boundaries are not optimized to give the best discrimination ability: HCA clusters based on a cascading nearest-neighbor method, while PCA develops principal components based on maximizing the variance among all points in the dataset. Rather than attempt to use these non-optimized methods to develop classifiers, we chose to use a common machine learning tool that was specifically developed to maximize discrimination ability: support vector machines (SVM).<sup>40</sup>

**Support vector machines.** Unlike unsupervised methods such as PCA, HCA, or other clustering methods, SVM is a predictive method that is designed to classify incoming data that is not part of the training database. SVM classification is based on pairwise class prediction and focuses on the data most likely to be misclassified (*i.e.*, data vectors near the decision boundary for any given class pair, the so-called support vectors) to create optimized decision boundaries that best separate the data for each given pair of classes in high dimensional space. The result of each pairwise comparison gives a vote that is used to determine the final classification.<sup>40</sup> A general graphical explanation of the process is shown in Fig. 6. SVM optimization factors have been fully developed and incorporated into LIBSVM, an open-source SVM library.<sup>41</sup>

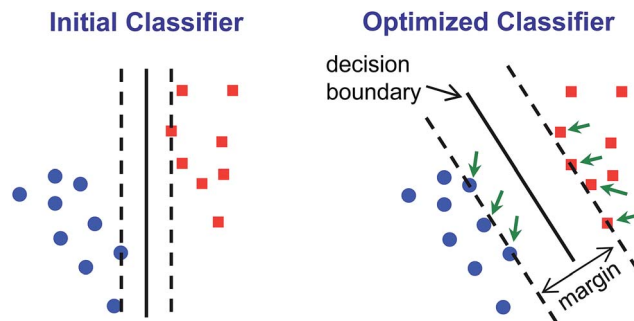


Fig. 6 Graphical illustration of SVM classifier optimization. A simplified initial guess is performed (left) and then algorithmically optimized through multiple iterations to maximize discrimination ability (typically, by maximizing the size of the margin and minimizing offside errors, as shown on the right). The margin is defined as the distance from contentious points (*i.e.*, support vectors, indicated by green arrows) to the decision boundary.

SVM is well optimized for discrimination within multidimensional datasets and has been widely and successfully developed, for example, for identification of objects in machine vision. Implementation of SVM uses multiple rounds of iteration to optimize parameters. Typical use involves data transformation using a kernel to convert the dataset into a linearly separable arrangement (*i.e.*, data lie on one side of a plane or the other, but not both). Using the colorimetric sensor arrays, the class data were found to be roughly normally distributed and linearly separable; no data transformation was necessary (*i.e.*, a linear kernel was used). Default SVM parameters were found to be quite well-optimized for the colorimetric sensor array database; this is not surprising, since HCA results already showed a high level of separation using a Euclidean distance clustering method.

Classifiers for each pair of analyte groups took the form of a decision hyperplane defined by an orthogonal 120-dimensional vector (*i.e.*, optimized linear combinations of  $\Delta$ RGB values of the sensor array) combined with a scalar value marking the position of the decision boundary; implementation in the automated handheld platform was simple, as it only required projection of the incoming 120-dimensional data vector from each scan onto the classifier vector using an inner product and comparison to the decision boundary scalar.

**Classification accuracy.** Classification accuracy of the SVM method was estimated using a leave-one-out cross-validation method. The database was divided into training sets and evaluation sets in a permutative manner: classifiers are created based on training sets (*i.e.*, with one trial left out) and predictions then made on the evaluation sets (*i.e.*, the left-out trial) and iterated among all permutations. Cross-validation results are shown in Table 2. For 12 of the analytes, no errors during cross-validation were observed among the septuplicate trials; the two  $\text{KClO}_3$  mixtures ( $\text{KClO}_3$ -sugar and  $\text{KClO}_3$ -fuel oil) and two nitro-organics (DMDNB and PETN), however, were non-separable within their respective groupings. In comparison to HCA, SVM was able to completely differentiate RDX (a nitroamine) from DMDNB and PETN (a nitroalkane and an alkyl nitrate ester, respectively).







- 11 C. McDonagh, C. S. Burke and B. D. MacCraith, *Chem. Rev.*, 2008, **108**, 400–422.
- 12 S. H. Lim, L. Feng, J. W. Kemling, C. J. Musto and K. S. Suslick, *Nat. Chem.*, 2009, **1**, 562–567.
- 13 M. C. Janzen, J. B. Ponder, D. P. Bailey, C. K. Ingison and K. S. Suslick, *Anal. Chem.*, 2006, **78**, 3591–3600.
- 14 K. S. Suslick, *MRS Bull.*, 2004, **29**, 720–725.
- 15 N. A. Rakow and K. S. Suslick, *Nature*, 2000, **406**, 710–713.
- 16 J. M. Rankin, Q. Zhang, M. K. LaGasse, Y. Zhang, J. R. Askim and K. S. Suslick, *Analyst*, 2015, **140**, 2613–2617.
- 17 L. Feng, C. J. Musto, J. W. Kemling, S. H. Lim, W. Zhong and K. S. Suslick, *Anal. Chem.*, 2010, **82**, 9433–9440.
- 18 H. Lin and K. S. Suslick, *J. Am. Chem. Soc.*, 2010, **132**, 15519–15521.
- 19 B. A. Suslick, L. Feng and K. S. Suslick, *Anal. Chem.*, 2010, **82**, 2067–2073.
- 20 Z. Li, M. Jang, J. R. Askim and K. S. Suslick, *Analyst*, 2015, **140**, 5929–5935.
- 21 Z. Li, W. P. Bassett, J. R. Askim and K. S. Suslick, *Chem. Commun.*, 2015, **51**, DOI: 10.1039/c1035cc06221g, in press.
- 22 J. M. Rankin and K. S. Suslick, *Chem. Commun.*, 2015, **51**, 8920–8923.
- 23 J. R. Carey, K. S. Suslick, K. I. Hulkower, J. A. Imlay, K. R. C. Imlay, C. K. Ingison, J. B. Ponder, A. Sen and A. E. Wittrig, *J. Am. Chem. Soc.*, 2011, **133**, 7571–7576.
- 24 C. L. Lonsdale, B. Taba, N. Queraltó, R. A. Lukaszewski, R. A. Martino, P. A. Rhodes and S. H. Lim, *PLoS One*, 2013, **8**, e62726.
- 25 Y. Zhang, J. R. Askim, W. Zhong, P. Orlean and K. S. Suslick, *Analyst*, 2014, **139**, 1922–1928.
- 26 J. R. Askim and K. S. Suslick, *Anal. Chem.*, 2015, **87**, 7810–7816.
- 27 A. García, M. M. Erenas, E. D. Marinetto, C. A. Abad, I. de Orbe-Paya, A. J. Palma and L. F. Capitán-Vallvey, *Sens. Actuators, B*, 2011, **156**, 350–359.
- 28 M. O. Salles, G. N. Meloni, W. R. de Araujo and T. R. L. C. Paixão, *Anal. Methods*, 2014, **6**, 2047–2052.
- 29 L. Shen, J. A. Hagen and I. Papautsky, *Lab Chip*, 2012, **12**, 4240–4243.
- 30 D.-S. Lee, B. G. Jeon, C. Ihm, J.-K. Park and M. Y. Jung, *Lab Chip*, 2011, **11**, 120–126.
- 31 L. Feng, C. J. Musto, J. W. Kemling, S. H. Lim and K. S. Suslick, *Chem. Commun.*, 2010, **46**, 2037–2039.
- 32 H. J. H. Fenton, *J. Chem. Soc., Trans.*, 1894, **65**, 899–910.
- 33 C. F. Poole, *Gas Chromatography*, Elsevier Science, Oxford, UK, 2012.
- 34 J. H. Robins, G. D. Abrams and J. A. Pincock, *Can. J. Chem.*, 1980, **58**, 339–347.
- 35 M. Crampton, *Adv. Phys. Org. Chem.*, 1969, **7**, 211–257.
- 36 J. F. Hair, B. Black, B. Babin, R. E. Anderson and R. L. Tatham, *Multivariate Data Analysis*, Prentice Hall, New York, 6th edn, 2005.
- 37 J. Janata, *Principles of Chemical Sensors*, Springer, New York, 2nd edn, 2009.
- 38 R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, Prentice Hall, New York, 6th edn, 2007.
- 39 S. Stewart, M. A. Ivy and E. V. Anslyn, *Chem. Soc. Rev.*, 2014, **43**, 70–84.
- 40 P. Flach, *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*, Cambridge University Press, 2012.
- 41 C.-C. Chang and C.-J. Lin, *ACM Trans. Intell. Syst. Technol.*, 2011, **2**, 1–27.
- 42 H. Lin, M. Jang and K. S. Suslick, *J. Am. Chem. Soc.*, 2011, **133**, 16786–16789.
- 43 *Convention of the Marking of Plastic Explosives for the Purpose of Detection*, International Civil Aviation Organization, Montreal, Canada, 1991, [https://www.unodc.org/tldb/en/1991\\_Convention\\_Plastic\\_Explosives.html](https://www.unodc.org/tldb/en/1991_Convention_Plastic_Explosives.html).
- 44 *Public Law 104–132: Antiterrorism and Effective Death Penalty Act of 1996*, US Government, Washington, DC, 1996, <http://www.gpo.gov/fdsys/pkg/PLAW-104publ132/pdf/PLAW-104publ132.pdf>.
- 45 *300-600 DPI Contact Image Sensor*, [http://www.csensor.com/M116\\_CIS.htm](http://www.csensor.com/M116_CIS.htm), accessed 28 January, 2014, 2014.
- 46 G. R. Peterson, W. P. Bassett, B. L. Weeks and L. J. Hope-Weeks, *Cryst. Growth Des.*, 2013, **13**, 2307–2311.
- 47 A. Wierzbicki, E. A. Salter, E. A. Cioffi and E. D. Stevens, *J. Phys. Chem. A*, 2001, **105**, 8763–8768.

