Analytical Methods

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Journal Name RSCPublishing

COMMUNICATION

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Fluorescence Sensor Array for Identification of Foodborne Pathogens

Radeemada Mungkarndee,[a] Gamolwan Tumcharern,[b] Raweewan Thiramanus,[b] Ittipon Techakriengkrai[c] and Mongkol Sukwattanasinitt*[d]

A fluorescence sensor array containing polyelectrolyte fluorophores is developed for foodborne pathogen detection and identification. The fluorophore-cell surface interactions generate fluorescence response patterns that can be differentiated by linear discriminant analysis (LDA) to identify 8 bacteria with the classification accuracy of 100%.

Safe and sufficient food is a fundamental human need. Foodborne diseases are a global health concern encompassing millions of people. A wide spectrum of illnesses is the result of ingestion of foodstuffs majorly contaminated by microbial pathogens. In the past decade, serious outbreaks of foodborne disease have been reported on every part of the world which not only adversely affect people's morbidity and mortality, but also have negative economic consequences for society, organizations, public and private, communities and individuals. The term "foodborne disease" has been traditionally defined as illnesses caused by microorganisms, with often acute reactions, such as diarrhoea. World Health Organization (WHO) estimates that worldwide foodborne and waterborne diarrheal diseases kill about 2.2 million people annually.[1] Conventional and standard bacterial detection methods such as culture and colony counting methods may take up to a few days to yield a result, while immunology-based methods and polymerase chain reaction based methods, require extensive sample preparation.[2,3] Optical sensors for biological agents are promising to overcome these limitations and become a feasible choice for a rapid detection in less complex samples such as drinking water. The concept behind this work is that our fluorescent compounds having various interaction sites differently bind onto the pathogenic bacteria cell surface by electrostatic, hydrophilic and hydrophobic interactions. Thus, the challenges of our sensor involve the recognition of cell structural complexity, including cell morphology and outer membrane of the target analytes.

We have previously reported the dendritic polyelectrolyte fluorescent compounds **1**-**3** (Fig. 1), having various interaction sites, created by different combinations of cationic trimethylammonium, anionic carboxylate and non-ionic methyl ester on the peripheries, for the application as a sensor array for discrimination of eight proteins.[4] In this study, we aim to investigate the discrimination ability of this set of fluorophores in the pathogen identification based on the variation of the bacterial cell surfaces.

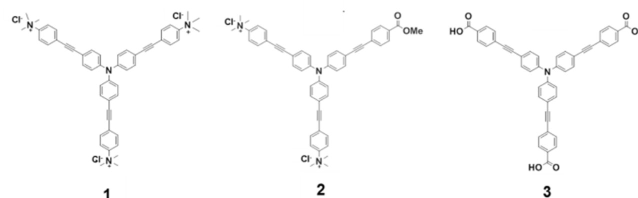

Fig. 1 Structures of fluorophores.

All fluorophore and bacteria solutions were prepared by using Milli-Q water (18.1 MΩ) PBS buffer (pH 7.4) as the solvent. All chemicals were reagent grade and used as received without further purification. Fluorescence spectra were acquired from a SpectraMax M2 microplate reader (Molecular Devices, Sunnyvale, CA) Varian Cary Eclipse spectrofluorometer using black polystyrene 96-well microplates. The number of bacterial counts (CFU/well) was estimated from the optical density (OD) at 600 nm, using a calibration line related to the plate counts. For the fluorescence measurement, the number of bacterial counts was controlled at 10^8 CFU/well. The stock solutions of all fluorophores were prepared in 10 mM sodium phosphate buffer saline (PBS) pH 7.4. For the fluorescence measurement, the protein and fluorophore mixture was prepared by mixing the corresponding stock solutions and diluting with PBS to afford the final fluorophore concentration of 0.7 µM and incubating for 15 minutes.

Fluorescent molecules as sensor array have been developed for microorganism detection via pattern recognition of fluorescence

Analytical Methods Accepted Manuscript

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

responses using multivariate statistical analyses such as principal component analysis (PCA) and linear discriminant analysis (LDA).[2,3] Eight bacteria i.e. *Vibrio cholera* (Gram-negative, curved rod shape), *Shigella flexneri* (Gram-negative, rod shape), *Bacillus cereus* (Gram-positive, rod shape), enterotoxigenic *Escherichia coli* and non-enterotoxigenic *Escherichia coli* (Gram-negative, rod shape), *Listeria monocytogenes* (Gram-positive, rod shape), *Salmonella typhi* (Gram-negative, rod shape) and *Staphylococcus aureus* (Gram-positive, round shape) were selected as the samples for testing the discrimination power of the fluorophores **1**-**3** sensor array. These bacteria are pathogenic microorganism responsible for various foodborne illnesses (Table S1).[4,7-9] The fluorescence responses of the fluorophores upon mixing with the solution containing pathogens (10^8 CFU/well) were measured in the range of 400-700 nm using the excitation wavelength at 375 nm in a fluorescence microplate. The intensity differences ΔI, calculated from $I_E – I_0$, where I_E and I_0 were the fluorescence intensity in the presence and absence of the bacteria, at all wavelength were collected and analysed by multivariate statistical analyses.[10,11] The fluorogenic responses (ΔI) obtained from the fluorescence measurements were prepared as data matrix of $p \times n$ where p (row) corresponds to the numbers of bacteria samples time numbers of repetitions and n (column) corresponds to the numbers of fluorophores time numbers of wavelengths.

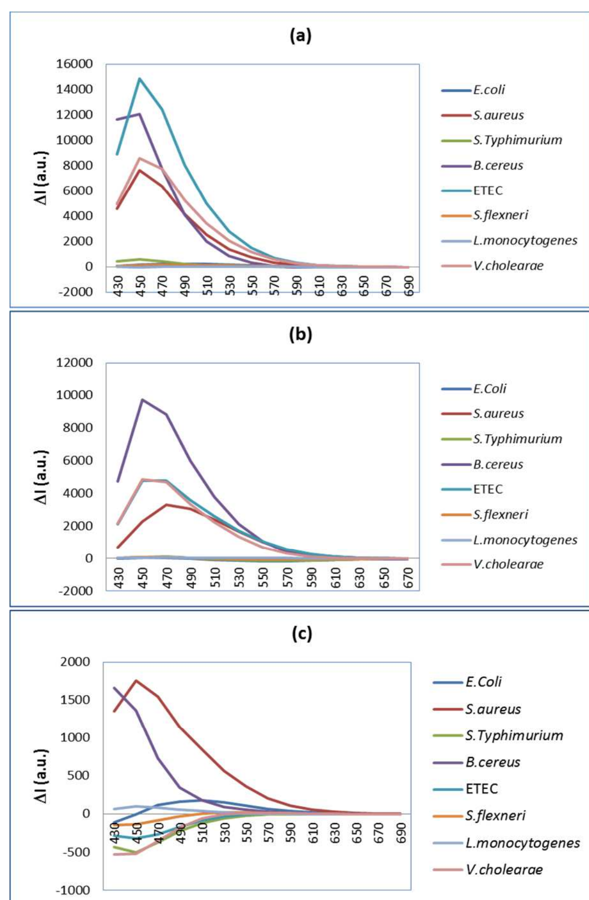

Fig. 2 Fluorescence responses (ΔI) of fluorophore (a) **1** (b) **2** and (c) **3** upon addition of the solution containing pathogens. Each spectral line is an averaged ΔI of 21 repetitive samples.

Fig. 2 shows that the fluorescence signal of **1** was enhanced by ETEC > *B.cereus* > *V.cholearae* > *S.aureus*, and almost no response to the rest of the bacteria. Fluorophore **2** showed fluorescence enhancement with *B.cereus* > ETEC ~ *V.cholearae* > *S.aureus*, and almost no response to the rest of the bacteria. Fluorophore **3** gave both fluorescence enhancement (*S.aureus* > *B.cereus* > *E.coli* > *L.monocytogenes*) and quenching (*V.cholearae* ≅ *S.Typhimurium* > ETEC > *S.flexneri*).

Principal component analysis (PCA) and linear discriminant analysis (LDA) were performed on the $p \times n$ data matrix using XLStat 2010. Full cross-validation with a leave-one-out technique was applied to both PCA and LDA models to assess the performance of each model based on the classification accuracy of the samples in the validation set. In the cross-validation, a sample in p was randomly removed from the data set and LDA was used to determine the centroid coordinate for each known class of the rest of the samples (p-1). The removed sample was then classified to the group of which centroid closest to the sample score coordinate. The procedure was repeated until all samples were classified. The ratio of the numbers of the samples correctly predicted by LDA to the total numbers of the samples defined the accuracy percentage.

PCA is an unsupervised statistical method which condense large amounts of data into fewer latent variables called principal components (PCs), while preserving intrinsic variance of the original data as much as possible.[12,13] The first PC contains the highest degree of data variance and other PCs follow in the order of decreasing variance. In this study, PCA was applied to convert the fluorescence dataset with 42 original variances (3 flurophores × 14 wavelengths) into PC scores of PC1 and PC2 which accounted for 97.7% of the total variance. The PC score plot (Fig. 3) of all 168 samples (8 bacterial samples × 21 repetitions) on PC1 and PC2 coordinates gave four distinctive clusters of *V. cholera*, *S. aureus*, ETEC and *B. cereus* distributing on the right half of the plot, while *S. flexneri*, *E. coli*, *L. monocytogenes* and *S. Typhimurium* gathered as one cluster on the left hand and cannot be discriminated. The use of 3 sensing fluorophores combined with PCA technique is not likely to recognize the difference of bacteria cell envelope and shape as 3 out of 4 bacteria were in the same cluster, *S. flexneri*, *E. coli* and *S. Typhimurium* are Gram-negative and rod shape. In Gram-negative bacteria, the surface of the outer membrane is composed predominantly of lipopolysaccharides,[14] amphiphatic molecules which could be bound with the fluorophores **1**-**3**. The only Gram-positive bacterium in the bunch is *L. monocytogenes* but it is different from other Gram-positive bacteria for possessing of lipopolysaccharide at the cell surface resemble to the Gram-negative bacteria.[15,16] ETEC scores are at the lower right quadrant of the PCA score plot distinctively separated from the E. coli cluster which is on the left side of the plot. ETEC is a Gram-negative rod-shaped bacterium that produces more enterotoxin causing secretion of large amounts of fluids and electrolytes that result in diarrhea.[14,17-19] It typically adhere to host cells via filamentous bacterial surface structures known as colonization factors (CFs) [20,21] which possesses Coli Surface Antigen (CS), an antigen located in the outer surface coat. CS expressing charged residues at cell surface makes ETEC fluorescent pattern distinguished from the less harmful *E. coli* species.

Analytical Methods Accepted Manuscript

 This journal is © The Royal Society of Chemistry 2012

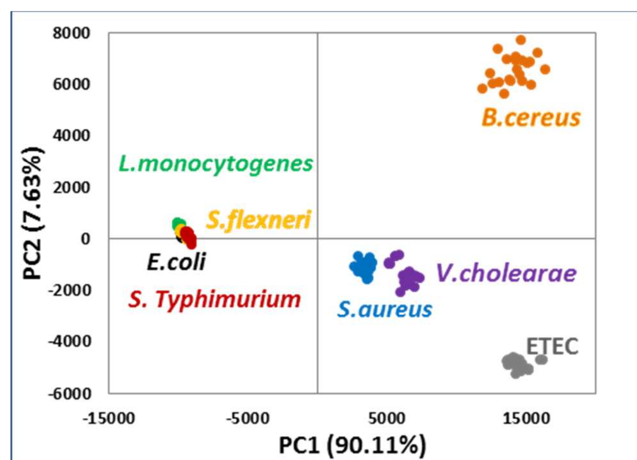

Fig. 3 PCA score plot of ΔI values of fluorophores **1**-**3** upon mixing with each bacteria sample.

Since PCA is the unsupervised method, the generated clusters of data are based on their similarities and differences perceived by the sensing elements without taking into consideration of analytes class labels.[13] Evidently, the results described above demonstrated that PCA did not have enough discrimination ability of some bacteria classes. The supervised LDA which normally gives superior discrimination ability than the PCA was then further investigated in our study. LDA is probably the most frequently used supervised pattern recognition method for complex samples such as in food analysis.[10] In order to achieve an optimum classification capacity of our sensor array, we applied the LDA analysis to process the same spectroscopic dataset. LDA is based on the determination of linear discriminant functions, which will find the directions (axes) that maximize the linear separation among the multiple groups of analytes.[10,22,23]

The fluorescence response patterns were subjected to linear discriminant analysis (LDA); which converts the data matrix of 7,056 ΔI values (3 flurophores × 14 wavelengths × 8 bacterial samples × 21 repetitions) to discriminant scores. The 3-D plot showed 8 well defined clusters corresponding to the bacterial pathogens (Fig. 4). The first three discriminant factors (F1, F2 and F3) contain 58.08, 27.69, and 8.69% respectively, occupying 94.46% of total variation. Fig. 4 shows 8 clusters with no overlap between the groups. The leave-one-out cross-validation of the LDA scores also revealed a classification accuracy of 100%.

Part of the motivation of this work was to practically apply our fluorescence sensor as a rapid detection method. We thus attempted to reduce sensing elements as minimal as possible by selecting a subset of fluorescence sensor array without deteriorating its discriminatory capacity.[24] In addition, the removal of irrelevant or noise variables should improve the performance of the sensor array because not every variable would be always significant for the discriminating ability. The LDA correlation plot (Fig. 5) was used to identify the importance of each individual fluorophore at each wavelength to each linear discriminant factor (F).

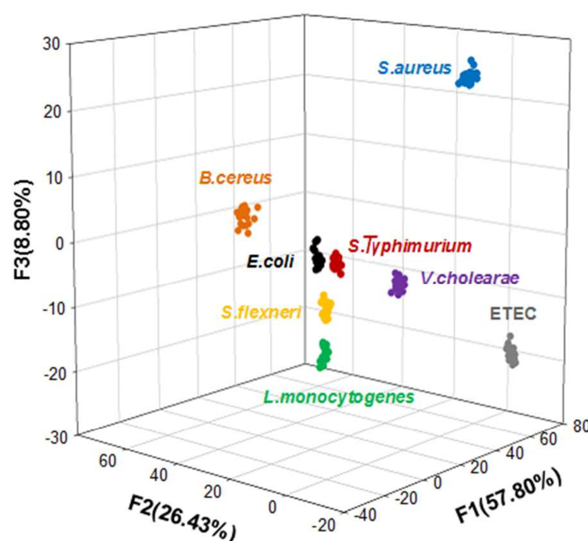

Fig. 4 LDA score plot of first three discriminant factors (F1, F2 and F3) obtained from fluorescence responses data of bacteria samples.

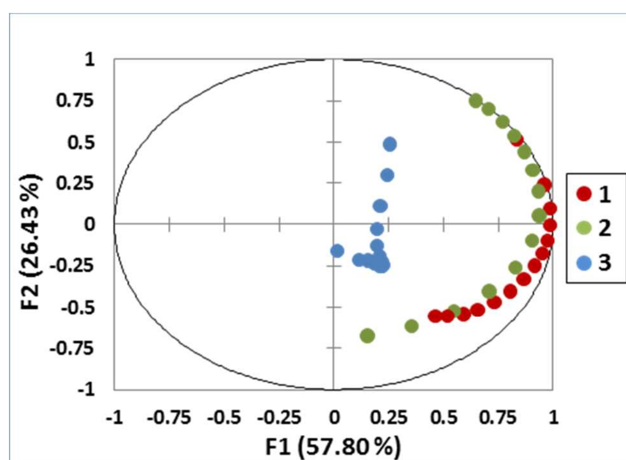

Fig. 5 LDA loading plot of ΔI values of fluorophores **1**-**3** upon mixing with each bacteria sample.

Fig. 5 shows that **3** yielded the lowest discriminant factor for both F1 and F2 while **1** and **2** yielded high discriminant factors for both F1 and F2. Compound **3** was thus removed from the array due to its low performance. The dataset without fluorogenic responses obtained from **3** was applied to LDA for further investigation.

Fig. 6 shows the LDA score plot for the LDA with cross validation performed for the sensor arrays of **1** and **2**. The first two Fs (F1 and F2) contained 85.76% of the variance and all 8 bacteria samples could be 100% accurately classified. The result confirms that the fluorescence dataset from **1** and **2** contains highly correlated variables which enable effective visualization and classification of multivariate data. Compound **3** can be removed from the sensor array without losing its discrimination power.

Analytical Methods Accepted Manuscript

This journal is © The Royal Society of Chemistry 2012

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

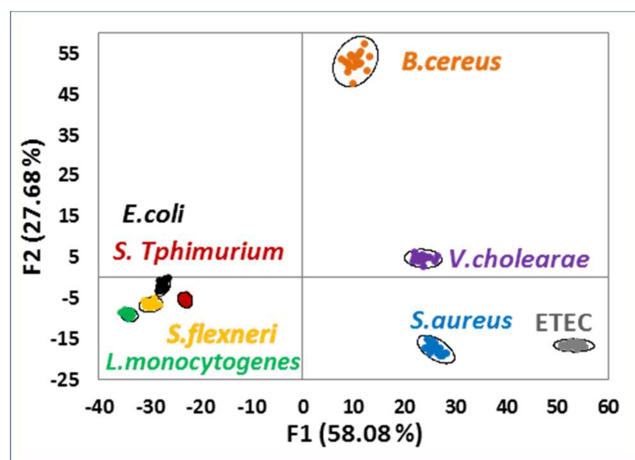

Fig. 6 LDA score plot of first two discriminant factors (F1 and F2) obtained from ΔI values of fluorophores **1** and **2** upon mixing with each bacteria sample. Oval outlines indicate groups of bacteria sample at 95% confidence level.

The test was also performed at a lower bacteria concentration of 10^3 CFU/well. The LDA score plot of the dataset obtained from 8 bacteria solutions × 9 repetitions × 2 fluorophores × 45 wavelengths (Fig. 7) accurately showed clusters of 8 bacteria on the first two discriminant factors containing 96.59% of the total variance. The cross-validation routine also gave a 100% level of correct classification for all samples.

To demonstrate discriminating applicability of our sensor array in real drinking water, two types of commercially available bottled drinking waters, mineral and non-mineral ones were tested. The mineral water samples were spiked with each of 7 pathogens (*V. cholera*, *S. aureus*, *B. cereus*, ETEC, *E.coli*, *L. monocytogenes* and *S. Typhimurium*) and the non-mineral water samples were spiked with each of 5 pathogens (*V. cholera*, *S. aureus*, *B. cereus*, ETEC and *E.coli*) at the concentration of 10^8 CFU/well. The mineral water samples gave different fluorescence responses from the non-mineral water samples (Fig. S15) indicating that the ions in the mineral water significantly altered the electrostatic interaction between the fluorophores and bacteria.[25] This effect may further complicate the data analysis that could reduce the discrimination ability of the sensor array. LDA was applied, without taking the water types into account, the dataset of 16,104 (12 bacteria-spiked waters × 11 repetitions × 2 fluorophores × 61 wavelengths) and processed to generate clusters of LDA scores corresponding to the numbers of bacteria types. The LDA score plot (Fig.8a) showed that F1 and F2 contained 76.51% of the variance and the cross-validation gave only 2 misclassified samples out of 132 samples (12 bacteria-spiked water × 11 repetitions) representing a classification accuracy of 98.48%. The LDA correlation plot (Fig. S16) demonstrated that fluorophore **1** gave well spread out loading values on F1, while fluorophore **2** provided relatively invariant effect on F1 suggesting that **1** contributed more significantly to the discriminatory performance of the array. To reduce the number of sensing elements, only **1** was selected to construct the array. The LDA score plot (Fig. 8b) gave no-overlapping clusters between the bacteria types and the leave-one-out cross-validation routine also confirmed 100% classification accuracy by fluorophore **1**. LDA was also performed on minimized array of **2** (Fig. 8c) but it gave discriminatory ability of only 96.97% accuracy. The results demonstrated that the identification of food pathogens in drinking water may be achieved by using just only fluorophore **1.**

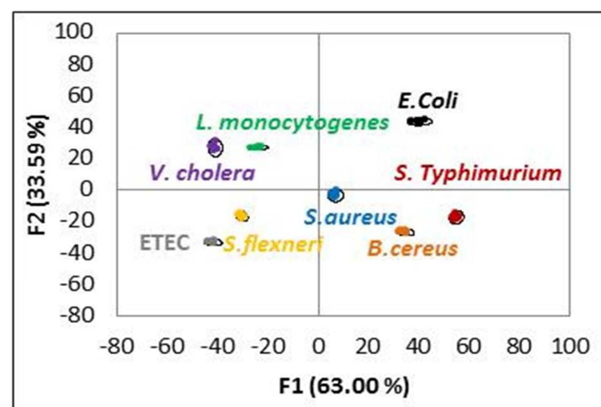

Fig. 7 LDA score plot of first two discriminant factors (F1 and F2) obtained from ΔI values of fluorophores **1** and **2** upon mixing with each bacteria–spiked drinking water at concentration of 10^3 CFU/well. Oval outlines indicate groups of bacteria sample at 95% confidence level.

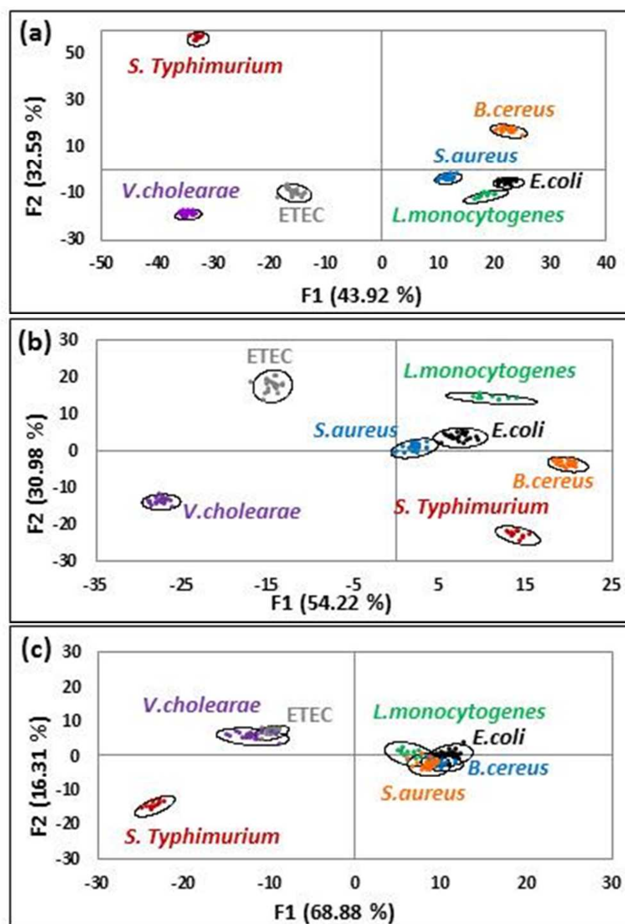

Fig. 8 LDA score plot of first two discriminant factors (F1 and F2) obtained from ΔI values of fluorophores (a) **1** and **2**, (b) **1** and (c) **2** upon mixing with each bacteria–spiked drinking water. Oval outlines indicate groups of bacteria sample at 95% confidence level.

Analytical Methods Accepted Manuscript

This journal is © The Royal Society of Chemistry 2012

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Conclusions

A fluorescence sensor array composed of three synthetic fluorophores has been investigated and successfully applied for discriminating bacterial pathogens. Statistical pattern recognition techniques, PCA and LDA, were applied on the fluorescence responses. Optimum classification result was achieved by LDA which was outperformed in recognizing the similarities within the group of bacteria samples. We demonstrated that fluorescent sensor array could identify a single type of foodborne pathogens in contaminated drinking water including the mineral water.

Acknowledgement

This study is financially supported by Nanotechnology Center (NANOTEC), NSTDA, Ministry of Science and Technology, Thailand, through its program of Center of Excellence Network. This work is part of the Project for Establishment of Comprehensive Center for Innovative Food, Health Products and Agriculture supported by the Thai Government Stimulus Package 2 (TKK2555, SP2) and the Ratchadaphiseksomphot Endowment Fund of Chulalongkorn University (RES560530126-AM).

Notes and references

[a] Program in Biotechnology, Faculty of Science, Chulalongkorn University, Phayathai Rd., Bangkok 10330, Thailand.
[b] National Nanotechnology Center, National Science and Technology Development Agency, Pathumthani 12120, Thailand.
[c] Department of Food Technology, Faculty of Science, Ramkhamhaeng University, Bangkok 10240, Thailand.
[d] Organic Synthesis Unit, Department of Chemistry, Faculty of Science and Nanotec-CU Center of Excellence on Food and Agriculture Chulalongkorn University, Bangkok 10330, Thailand. E-mail: smongkol@chula.ac.th; Tel.: +66 (2) 2187620; Fax.: +66 (2) 2187598

† Footnotes should appear here. These might include comments relevant to but not central to the matter under discussion, limited experimental and spectral data, and crystallographic data.

1 World Health Organization, (http://apps.who.int/iris/bitstream/10665/101542/1/9789241506281_eng.pdf) (accessed October 2014)
2 R. L. Phillips, O. R. Miranda, C.-C. You, V. M. Rotello, and U. H. F. Bunz, *Angew. Chem. Int. Ed.,* 2008, **47**, 2590 –2594.
3. Y. Wan, Y. Sun, P. Qi, P. Wang, and D. Zhang, *Biosens. Bioelectron.*, 2014. **55**, 289-293.
4 V. Velusamy, K. Arshak, O. Korostynska, K Oliwa and C. Adley, *Biotechnol. Adv.*, 2010, **28**, 232–254.
5 A. Singh, S. Poshtiban and S. Evoy, *Sensors*, 2013, **13**, 1763-1786.
6 N. Niamnont, R. Mungkarndee, I. Techakriengkrai, P. Rashatasakhon and M. Sukwattanasinitt, *Biosens. Bioelectron.*, 2010, **26**, 863–867.
7 Food and Drug Administration, (http://www.fda.gov/downloads/Food/FoodSafety/FoodborneIllness/FoodborneIllnessFoodbornePathogensNaturalToxins/BadBugBook/UCM297627.pdf) (accessed October 2014)
8 Washington State Department of Health (http://www.doh.wa.gov/Portals/1/Documents/5100/foodchart.pdf) (accessed October 2014)
9 Food Standards Australia New Zealand (http://www.foodstandards.gov.au/publications/Documents/FSANZ_FoodborneIllness_2013_WEB.pdf) (accessed October 2014)
10 L. Berrueta, A., R. M. Alonso-Salces and K´. H´eberger, *J. Chromatogr. A*, 2007, **1158**, 196–214.
11 L. Mutihac and R. Mutihac, *Anal. Chim. Acta,* 2008, **612**, 1–18.
12 E. Dufour, A. Letort, A. Laguet, A. Lebecque and J.N. Serra, *Anal. Chim. Acta,* 2006, **563**, 292-299.
13 M. A. Palacios, Z. Wang, V. A. Montes, G.V. Zyryanov and P. Anzenbacher, *J. Am. Chem. Soc.*, 2008, **130**, 10307–10314.
14 G. J. Tortora, B. R. Funke and C. L. Case in *Microbiology: An Introduction*, Benjamin Cummings, San Francisco, 10th edn., 2010, ch. 4, pp 84-87.
15 J. M. Farber and P. I. Peterkin, *Microbiol. Rev.*, 1991. **55**, 476-511.
16 F. Fiedler, *Infection*, 1988, **16**, S92– S97.
17 H. W. Moon, R. E. Isaacson and J. Pohlenz, *Am. J. Clin. Nutr.*, 1979, **32**, 119-127.
18 A. Cravioto, S. M. Scotland and B. Rowe, *Infect. Immun.*, 1982, **36**, 189-197.
19 A.-M. Svennerholm, C. Wenneras, J. Holmgren, M. M. McConnell and B. Rowe, *Infect. Immun.*, 1990, **58**, 341-346.
20 C. S. Alves, M. N. Melo, H. G. Franquelim, R., Ferre, M. Planas, L. Feliu, E. Bardají, W. Kowalczyk, D. Andreu, N. C. Santos, M. X. Fernandes and M. A. R. B. Castanho, *J. Biol Chem*, 2010, **285**, 27536–27544.
21 *US Pat.*, 2006 0153878 A1, 2006
22 C. C. You, O. R. Miranda, B. Gider, P. S. Ghosh, I.-B. Kim, B. Erdogan, S. A. Krovi, U. H. F. Bunz and V. M. Rotello, *Nat. Nanotechnol.*, 2007, **2**, 318-323.
23 S. F. Vallejuelo, G. Arana, A. Diego and J, M. Madariaga, *Chemosphere,* 2011, **85**, 1347–1352.
24 J. Hixon and Y. K. Reshetnyak, *Algorithms*, 2009, **2**, 1155-1176.
25 N. Niamnont, W. Siripornnoppakhun, P. Rashatasakhon and M. Sukwattanasinitt, *Org. Lett.*, 2009, **11**, 2768–2771.

Analytical Methods Accepted Manuscript