Note

Special Issue: Cell and Biomolecule Analysis

Array-based generation of response patterns with common fluorescent dyes for identification of proteins and cells.

Masahiro Mimura,¹ Shunsuke Tomita,²* Ryoji Kurita,¹,² and Kentaro Shiraki¹

¹ Faculty of Pure and Applied Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan
² Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST) and DAILAB, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan

* Corresponding author. Tel: +81-29-861-2634; Fax: +81-29-861-6177.

E-mail address: s.tomita@aist.go.jp
Abstract

A differential array consisting of commercially available common fluorescent dyes was constructed for the identification of proteins and human cancer cells. Fluorescence of dyes was differently altered by mixing with proteins and human cancer cells, generating response patterns that are unique to the analytes. Linear discriminant analysis of the obtained patterns enabled the accurate identification of eight proteins and three human cancer cells. As this system can be easily prepared, it would offer a unique opportunity for array-based differential biosensing.

Keywords: biosensing, cancer cells, fluorescent dyes, multivariate analysis, proteins
Introduction

Rapid and efficient identification of proteins and cells is of critical importance for biomedical research and disease diagnosis.\(^1\)\(^-\)\(^3\) In recent years, the use of array-based differential sensing for this purpose has increasingly attracted great attention.\(^4\)\(^,\)\(^5\) Unlike conventional “lock-and-key” specific recognition, array-based differential sensing employs multiple molecules capable of non-specifically interacting with target analytes. The addition of analytes to an array of such molecules results in the generation of response patterns. The generated patterns are then statistically processed with multivariate analyses to identify analytes.

Over the past decade, array-based differential sensing has allowed identifying various biological analytes, ranging from sugars,\(^6\)\(^-\)\(^8\) amino acids,\(^9\)\(^,\)\(^10\) proteins,\(^11\)\(^-\)\(^14\) bacteria\(^15\)\(^,\)\(^16\) to cells.\(^17\)\(^-\)\(^19\) For many cases of large analytes including proteins and cells, arrays of complexes consisting of multiple tailored materials are chosen in order to produce diverse and strong multipoint interactions. For example, Rotello et al. proposed the effectiveness of using complexes between gold-nanoparticles and fluorescent macromolecules, such as synthetic polymers and fluorescent proteins.\(^20\)\(^,\)\(^21\) Our group has also developed arrays of different types of complexes for the identification of proteins\(^22\)\(^-\)\(^25\) and cells.\(^26\)\(^,\)\(^27\)

An easy-to-prepare system for array-based differential sensing that does not require labor synthesis is crucial for practical applications.\(^28\) Here, we focused on commercially available common fluorescent dyes. Various degrees of fluorescence quenching or enhancing of dyes are caused by binding of proteins, depending on the characteristics of both dyes and proteins.\(^29\)\(^,\)\(^30\) It would be possible that fluorescent dyes with these properties can be used as sensor elements of a differential array. Based on this assumption, we have constructed an array-based differential sensing system consisting of common fluorescent dyes for accurate identification of proteins and human cancer cells.
Results and Discussion

The usefulness of fluorescent dyes to construct the easy-to-prepare system was investigated by using five selected commercially available fluorescent dyes with different charges; 8-methoxypyrene-1,3,6-trisulfonic acid (M$_3^-$), Eosin Y (E$_2^-$), fluorescein (F$^-$), acridine orange (A), rhodamine 6G (R$^+$) (Fig. 1A and Table S1). As individual fluorescent dyes are likely to possess different affinities toward proteins or surface components of cells, we hypothesized that the addition of analytes, such as proteins or cells, to an array of solutions containing fluorescent dyes would generate unique response patterns of changes in fluorescence intensity, which possibly provide multidimensional information of sufficient quality for discrimination by multivariate analyses (Fig. 1B).

In order to examine whether the fluorescence of the dyes is modulated by proteins, fluorescence titration of HSA to M$_3^-$ was initially carried out. Although negatively charged HSA (Fig. 1C) and M$_3^-$ were expected to be repulsed electrostatically to each other, the fluorescence emission of M$_3^-$ was decreased with increasing the concentration of HSA (Fig. 2A). The interaction of M$_3^-$ probably was related to the more local circumstances of HSA. Positively charged Cyt-C quenched fluorescence of M$_3^-$ up to almost 60% similarly to HSA, but Cyt-C caused fluorescence reduction more gently (Fig. 2B). The binding affinity of these proteins toward M$_3^-$ was estimated according to the literature; HSA and Cyt-C possess 1.16 and 3.07 binding sites governed by a dissociation constant ($K_{SV}$) = 1.69×10$^6$(M$^{-1}$) and 1.34×10$^5$(M$^{-1}$), respectively (For details, see Supporting Information). These results suggested that fluorescent dyes appeared to be able to provide the responses reflecting unique protein properties.

After determination of the appropriate fluorescent dye concentrations, we tested the applicability of the fluorescent dye-array to protein identification. Eight proteins with different molecular weight (Mw) and isoelectric point (pI) were chosen as sensing targets (Fig. 1C). For sensing procedure, each protein solution was added to each well of a 96-well plate, which contained solutions of fluorescent dyes in 20 mM 3-(N-morpholino)propanesulfonic acid (MOPS).
buffer (pH = 7.0), to reach a final concentration of 50 µg/mL proteins and 500 nM fluorescent dyes. The fluorescence intensities from individual wells were recorded with a microplate reader as \( (I - I_0) \) at two different channels (for detailed, see Materials and method in the Supporting Information), generating a data matrix of 5 fluorescent dyes × 2 channels × 8 proteins × 6 replicates (Table S2).

Anionic \( \text{M}^3- \), \( \text{E}^2- \) and \( \text{F}^- \) were quenched by all the proteins, while the fluorescence of neutral \( \text{A} \) and cationic \( \text{R}^+ \) was changed in various ways depending on the proteins (Fig. 3A). This suggested that the charge state of fluorescent dyes is associated with fluorescence changes. The resulting data matrix was then analyzed with linear discriminant analysis (LDA), a well-established multivariate analysis method for multigroup classification. A linear discriminant score plot revealed eight clusters corresponding to the individual proteins with slight overlap between Lip and IgG (Fig. 3B). The clusters of serum albumins were well-separated from other proteins, especially on the x-axes, which accounts for 93.1% of the total variance. Nevertheless, the residual five proteins showed a separation not only on the x-axis but also on the y-axis (3.7% of the total variance). Thus, it is possible that two or more characteristics of proteins were responsible for the generated responses. According to a leave-one-out cross-validation analysis, the so-called the Jackknifed classification procedure, the use of single fluorescent dyes afforded classification accuracies of 46% to 75%, while 100% accuracy was achieved when using five fluorescent dyes, indicating that the individual dyes possess different affinities for the proteins.

As indicated in the above and previous report, electrostatic interactions should play a significant role in the interactions between dyes and proteins. However, the first discriminant scores (Fig. 3B) showed a low correlation with the pIs of the proteins \( (r = -0.21) \), suggesting that other characteristics, especially hydrophobicity, are probably related. For example, binding of \( \text{E}^2- \) to HSA is caused by both electrostatic and hydrophobic interactions. As fluorescent dyes used in this study possess different degrees of hydrophobicity, with Log \( P \) values that vary from 0.48 (\( \text{M}^3- \)) to 6.05 (\( \text{E}^2- \)) (Fig. 1A), hydrophobic interactions would also contribute to generations of
diverse responses. In addition, specific interactions may be occurred, e.g., F\(^-\) specifically binds to a certain subdomain of HSA.\(^{35}\) Charge transfer occurring between bound dyes and proteins possibly be responsible for the quenching mechanism, as it has been proposed that fluorescence changes of xanthene dyes are associated with charge transfer to protein components, such as aromatic amino acids.\(^{29}\) Taken together, the high accuracy of our system should be attributed to these diverse interactions of fluorescent dyes toward proteins.

The discrimination power of the fluorescent dye-array was further examined using human cancer cells, as the detection of cancer cells is crucial for diagnosis and effective therapy.\(^{36,37}\) We selected three human cancer cell lines: MG63 (bone), A431 (skin) and HL60 (blood) (Fig. 1D). 2.5 × 10\(^4\) cells/mL of these cells were analyzed by the array (Table S3) and then processed by LDA. A linear discriminant score plot showed a clear separation of the clusters in a two-dimensional space (Fig. 4) with a classification accuracy of 100% in the Jackknifed classification procedure. These results suggested the versatility of our sensing system for the characterization of biological analytes. To differentiate more cells, it is necessary to obtain further diverse response patterns by, e.g., selection of functional dyes or detection of different optical properties of dyes.

The array consisting of common fluorescent dyes exhibits good discriminatory power, but it shows substantially lower sensitivity compared to previously reported our enzyme/polymer complexes-array that can amplifies the interactions with analyte proteins (\(~0.5\) μg/mL).\(^{27}\) In addition, unlike polymers capable of strongly interacting with proteins through multiple electrostatic interactions,\(^{26}\) it is presumed that low-molecular-weight common fluorescent dyes are interfered by contaminants included in biological samples, such as serum and cell culture media. Therefore, development of appropriate fluorescent dyes that address these issues would improve sensitivity and applicability of the system.
Conclusion

In summary, we have shown that even the commercially available common fluorescent dyes can be used to generate response patterns of proteins and human cancer cells. This easy-to-prepare system would be capable of being tuned and extended as other fluorescent dyes are also commercially available, hence the approach presented herein provides a unique opportunity for the construction of array-based differential biosensing system.

Acknowledgment

We would like to thank S. Ishihara for the technical assistance with the experiments. This work was supported by JSPS KAKENHI Grants JP16K14043 and JP17H04884.

Supporting Information

Experimental details and response profiles. This material is available free of charge on the Web at http://www.jsac.or.jp/analsci/.
References


Figure Captions

Fig. 1 (A) Chemical structure of fluorescent dyes used in this study. (B) Schematic illustration of the array-based differential sensing. (C, D) Profiles of (C) proteins and (D) human cancer cells used as analytes in this study. Log P values of dyes shown in parentheses were calculated using the program ALOGPs.38

Fig. 2 Changes in fluorescence intensity of M3+ upon addition of proteins. (A) Fluorescence spectra of 500 nM M3+ upon addition of HSA in 20 mM MOPS (pH = 7.0); λex =375 nm, [HSA] = 0-101 µg/mL. (B) fluorescence intensity of 500 nM M3+ in the presence of different concentration of Cyt-C and HSA in 20 mM MOPS (pH = 7.0); λex/λem = 375 nm/430 nm.

Fig. 3 Protein identification using a differential array consisting of fluorescent dyes. (A) Patterns of changes in the fluorescence intensity upon addition of protein solutions (50 µg/mL) from an array consisting of 500 nM of five fluorescent dyes. Values are shown as mean values ± SD (n = 6). (B) Discriminant score plot for proteins (50 µg/mL) obtained from the system of 5 fluorescent dyes × 2 channels, whereby ellipsoids represent confidence intervals (±1 SD) for the individual analytes.

Fig. 4 Discriminant score plot for human cancer cells (2.5 × 10⁴ cells/mL) obtained from the system of 5 fluorescent dyes × 2 channels, whereby ellipsoids represent confidence intervals (±1 SD) for the individual analytes.
Fig. 1
Fig. 2

Fig. 3
Fig. 4
Graphical Index

Commercially available common fluorescent dyes

Protein samples

Multivariate analysis

Protein identification with 100% accuracy

Score(1), 93.1%

Score(2), 3.7%