Fingerprint-based Protein Identification in Cell Culture Medium Using Environment-sensitive Turn-on Fluorescent Polymer

Hiroka Sugai,¹ Shunsuke Tomita,¹* Sayaka Ishihara,¹ and Ryoji Kurita¹,2**

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

(Received June 25, 2018; accepted August 24, 2018)

Keywords: proteins, biosensors, multivariate analysis, polymers, cell culture medium

The identification of secreted proteins in cell culture supernatants is a useful method for the noninvasive evaluation of cultured cells. Herein, we show that a fingerprint-based sensor technique can be used to identify typical hepatocyte-derived secretory proteins spiked into a cell culture medium. A poly-L-lysine modified with environment-sensitive dansyl groups (PLL-Dnc), which allows a turn-on fluorescent response with cross-reactivity against different analytes, was employed for the sensing of a series of secretory proteins (albumin, α₁-antitrypsin, fibrinogen, transferrin, and α-fetoprotein). An array of PLL-Dnc in different buffer solutions successfully produced fluorescence fingerprints as a result of distinct interactions with analyte proteins depending on solution conditions (pH and ionic strength), enabling the qualitative identification of five secretory proteins in culture media (40 μg/mL) with 100% accuracy using linear discriminant analysis. The array system was also capable of analyzing culture media that contain different concentrations of albumin and α-fetoprotein under realistic conditions. This work demonstrates the solution-condition-dependent discriminatory response of PLL-Dnc toward proteins spiked into a culture medium, rendering such a PLL-Dnc system a promising platform for the antibody-free and marker-based evaluation of cultured cells.

1. Introduction

Cells secrete diverse proteins to communicate with their surroundings. As these proteins reflect the state and function of cells for a given environment and time, the set of cell secretory proteins is regarded as a rich source of biomarkers for disease diagnosis and studies on drug sensitivity, as well as therapeutic outcome and prognosis.¹,² In the field of cell engineering, the biomarker proteins secreted into culture supernatants are routinely monitored for the control of cultured-cell quality, i.e., guarantee of cell activity and function³ and the in vitro assessment of chemotherapeutic agents.⁴,⁵ Antibody-based techniques are the most widely used for the identification of secretory proteins,⁶ with which the specific recognition of target proteins by

¹Corresponding author: e-mail: s.tomita@aist.go.jp
**Corresponding author: e-mail: r.kurita@aist.go.jp
https://doi.org/10.18494/SAM.2019.2032
ISSN 0914-4935 © MYU K.K.
https://myukk.org/
antibodies affords sensitive detection. However, this strategy presents some limitations, such as high production costs, structural instability, and difficulty with respect to the procurement of antibodies that are highly specific to target proteins.

Fingerprint-based sensing techniques provide a promising alternative approach to protein sensing.\(^7,^8\) In contrast to the specific recognition of antibodies, this approach is based on “cross-reactive” interactions of a probe array against analyte proteins. Such probe arrays are constructed to exhibit the chemical diversity necessary to respond distinctively to a variety of analyte proteins, allowing the generation of target-specific response fingerprints. This strategy has been used to identify proteins dissolved in buffer solutions,\(^9\text{–}^{12}\) serum,\(^13\text{–}^{16}\) and urine.\(^17,^{18}\) Most of the previously reported fingerprint-based sensors, however, require multiple synthetic probes as sensor elements for the construction of an array with reliable discrimination ability,\(^7,^8\) which involves laborious synthetic efforts.

Recently, we have developed a fingerprint-based sensing strategy that allows the discrimination of proteins using an array consisting of one synthetic polymer, namely, poly-L-lysine modified with environment-sensitive dansyl groups (PLL-Dnc).\(^19\) This strategy relies on the unique properties of PLL-Dnc, which affords turn-on fluorescence signaling with buffer-condition-dependent cross-reactivity against different proteins. The PLL-Dnc system can distinguish structurally similar albumins with/without post-translational modifications. Moving forward to the application of this system in the sensing of proteins secreted by cultured cells, we have examined the ability of PLL-Dnc to identify hepatocyte-derived secretory proteins spiked into culture media.

2. Materials and Methods

2.1 Materials

Albumin (ALB) from human serum, α\(_1\)-antitrypsin (AAT) from human plasma, fibrinogen (FIB) from human plasma, apo-transferrin (TF) from human, 2-morpholinoethanesulfonic acid (MES), 3-morpholinopropanesulfonic acid (MOPS), and L-glutamine were purchased from Sigma-Aldrich, Co., LLC (St. Louis, MO, USA). α-Fetoprotein (AFP) from human cord serum was obtained from Lee Biosolutions, Inc. (Maryland Heights, MO, USA). 3-[4-(2-Hydroxyethyl)-1-piperazinyl]propanesulfonic acid (EPPS) was purchased from Dojindo Molecular Technologies, Inc. (Kumamoto, Japan). Sodium chloride was obtained from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). 1X CD CHO medium was purchased from Thermo Fisher Scientific, Inc. (Waltham, MA, USA). Amicon Ultra 0.5 mL centrifugal filters with MWCO 30 kDa were obtained from Merck KGaA (Darmstadt, Germany). The PLL-Dnc polymer was synthesized as described in our previous report.\(^19\)

2.2 General procedures

CDCHO\(^+\) was prepared by supplementing the serum-free 1X CD CHO medium with L-glutamine to a final concentration of 8 mM. The protein concentration was spectroscopically determined by measuring the absorbance at 280 nm using a NanoDrop ND-
1000 spectrophotometer (Thermo Fisher Scientific, Inc., Waltham, MA, USA). Extinction coefficients at 280 nm ($\varepsilon_{280}$) were calculated by the method of Pace et al.$^{(20)}$ ALB, 0.518 (mg/mL)$^{-1}$ cm$^{-1}$; AAT, 0.450 (mg/mL)$^{-1}$ cm$^{-1}$; FIB, 1.689 (mg/mL)$^{-1}$ cm$^{-1}$; TF, 1.132 (mg/mL)$^{-1}$ cm$^{-1}$. For AFP, the $\varepsilon_{280}$ value provided by the manufacturer [0.520 (mg/mL)$^{-1}$ cm$^{-1}$] was used. The supplied solution of AFP was replaced with CDCHO+ using Amicon Ultra filters following the manufacturer’s instructions prior to the determination of the concentration. Fluorescence spectra and intensities were recorded on a Cytation5 Imaging Reader (BioTek Instruments, Inc., Winooski, VT, USA). The solutions were dispensed into Corning 96-well NBS microplates (Corning, NY, USA) using a PIPETMAX liquid handling system (Gilson, Inc., Middleton, WI, USA).

2.3 Fluorescence spectroscopic measurements

To each well of a 96-well microplate, we added 200 µL of a mixture of 5.0 µg/mL PLL-Dnc, 0–80 µg/mL ALB, 10% CDCHO+, and a buffer solution [18 mM MES (pH 5.5); 18 mM MES (pH 5.5) with 25 mM NaCl; 18 mM EPPS (pH 8.5)]. After incubation at 35 ℃ for 10 min, fluorescence spectra were recorded at an excitation wavelength ($\lambda_{ex}$) of 340 nm and emission wavelengths ($\lambda_{em}$) ranging from 380 to 650 nm.

2.4 Cross-reactive sensing and statistical analysis

To each well of a 96-well microplate, we added 160 µL of a buffer solution [22.5 mM MES (pH 5.5), 22.5 mM MOPS (pH 7.5), and 22.5 mM EPPS (pH 8.5) with/without 31.3 mM NaCl] followed by the addition of 20 µL of 50 µg/mL PLL-Dnc in deionized water. After incubation at 35 ℃ for 10 min, the fluorescence intensity was measured in two channels [$\lambda_{ex}$/$\lambda_{em}$: 340/480 (Ch1); 340/520 (Ch2)]. Then, 20 µL of 400 µg/mL protein in CDCHO+ was added to each well. After 10 min of incubation at 35 ℃, fluorescence was measured. The above process was repeated six times. Fluorescence responses are presented as $F - F_0$, where $F$ and $F_0$ are the fluorescence intensities after and before the addition of proteins, respectively. The obtained six datasets of fluorescence response were subjected to the classical linear discriminant analysis (LDA) in the SYSTAT software version 13 (Systat Inc., San Jose, CA, USA) to construct a training data matrix. For the blind test, the measurement of the fluorescence responses of the PLL-Dnc system was additionally repeated four times to generate a test data matrix. The obtained test data matrix was again subjected to LDA to classify test data into clusters constructed from the training data matrix on the basis of their smallest Mahalanobis distances.

3. Results and Discussion

3.1 Fluorescence responses of PLL-Dnc with albumin in the presence of culture medium

In recent years, the use of serum-free and chemically defined culture media has been required in cell engineering and biomedical fields for quality assurance and to avoid batch-to-batch variations and unnecessary suffering of unborn calves.$^{(21–23)}$ Therefore, a serum-
free and chemically defined culture medium (CDCHO+) was selected as the model medium in this study. The constructed sensing system is schematically illustrated in Fig. 1. On the basis of our previous work,\(^{(19)}\) we employed a polymer in which 4 out of 39 units of lysine side chains were modified with 5-(dimethylamino) naphthalene-1-sulfonyl, the so-called dansyl group [Fig. 1(a)]. When more hydrophobic dansyl groups are introduced into PLL, intra/intermolecular interactions between the dansyl groups increase the background fluorescence of the resulting PLL-Dnc. Therefore, the number of introduced dansyl groups was chosen to produce appropriate fluorescence responses with low background fluorescence to enable effective sensing. PLL-Dnc possesses positively charged ε-amino groups (\(p\)K\(_a\) ~10.5) and hydrophobic dansyl groups, which are capable of binding to various sites of the proteins nonspecifically via multiple electrostatic and hydrophobic interactions. When dansyl groups of PLL-Dnc bind to the hydrophobic surface region and the cavity of the proteins, they generate turn-on fluorescence signals. The predominant interactions between PLL-Dnc and the proteins depend on various conditions such as pH and ion strength; therefore, the use of different buffer solutions enables PLL-Dnc to produce the discriminatory fluorescence responses upon binding with the

Fig. 1. (Color online) Schematic representation of fingerprint-based protein identification using an environment-sensitive turn-on fluorescent polymer. (a) Chemical structure of PLL-Dnc. (b) Generation of fingerprints from turn-on fluorescence responses based on a dependence on the solution environment. (c) Fingerprint-based sensor array for protein identification.
analyte protein spiked into CDCHO+ [Fig. 1(b)]. Using different buffer solutions with PLL-Dnc as sensor elements, the fluorescence response fingerprints against each protein were produced and then processed by multivariate analysis to identify the corresponding proteins [Fig. 1(c)].

To examine whether the fluorescence responses of PLL-Dnc with proteins vary with the solvent, even in the presence of CDCHO+, we initially measured the fluorescence spectra of PLL-Dnc (5.0 μg/mL) solutions containing different concentrations of ALB in three different buffers [18 mM MES (pH 5.0); 18 mM MES (pH 5.0) with 25 mM NaCl; and 18 mM EPPS (pH 8.5)] with 10% CDCHO+. ALB was chosen as a representative secretory protein from hepatocytes.(24) As shown in Figs. 2(a)–2(c), clear hypsochromic shift and emission enhancement were observed for PLL-Dnc upon increasing the ALB concentration, which is consistent with the general response resulting from the interactions of the dansyl derivatives with the hydrophobic binding sites of proteins.(25) The maximum fluorescence intensities of PLL-Dnc toward ALB markedly varied depending on the buffer conditions, even in the presence of substantial amounts of

Fig. 2. (Color online) Turn-on fluorescence responses of PLL-Dnc (5.0 μg/mL) upon addition of ALB (0, 10, 20, 30, 40, 50, 60, 70, and 80 μg/mL) in three different buffer solutions with 10% CDCHO+. (a–c) Fluorescence spectra excited at 340 nm. The buffer solutions consisted of (a) 18 mM MES (pH 5.5), (b) 18 mM MES (pH 5.5) with 25 mM NaCl, and (c) 18 mM EPPS (pH 8.5). All spectra are the average of three scans. (d) Binding isotherms were monitored at 520 nm. Error bars represent the standard deviation (SD) from three values.
CDCHO+ [Fig. 2(d)]. The distinct responses should arise from differences in predominant factors (e.g., electrostatic interactions, hydrophobic interactions, surface heterogeneity, and steric hindrance) under respective conditions. From these results, we proceeded to determine the desired analyze proteins in subsequent sensing experiments at a concentration of 40 μg/mL, which was chosen as the optimum concentration for obtaining distinct fluorescence responses.

3.2 Sensing of hepatocyte-derived secretory proteins in the presence of culture medium

Once we confirmed that different buffer conditions afford distinct fluorescence responses, the PLL-Dnc system was applied to the fingerprint-based identification of proteins spiked into a culture medium. We selected five types of hepatocyte-derived secretory proteins as analyze proteins (Table 1). All these proteins are secreted from hepatocellular carcinoma-derived cell lines and some of them are employed as biomarkers to characterize cultured cells.\(^{(24)}\)

The model proteins in CDCHO+ (20 μL) were mixed with PLL-Dnc in six different buffer solutions (180 μL) with final concentrations of 5.0 μg/mL PLL-Dnc and 40 μg/mL analyze protein in the buffer solution [18 mM MES (pH 5.5), 18 mM MOPS (pH 7.0), and 18 mM EPPS (pH 8.0) with/without 25 mM NaCl]. The fluorescence responses were recorded in two different channels \([\lambda_{\text{ex}}/\lambda_{\text{em}}]: 340/480 \text{ (Ch1)}; 340/520 \text{ (Ch2)}\], generating a training data matrix of fluorescence responses (6 analytes × 6 solutions × 2 channels × 6 replicates) including a blank solution. The obtained fluorescence responses are visually summarized as a heat map in Fig. 3(a), where the color range from green to red denotes increasing fluorescence intensity under different buffer conditions. Despite slight differences in the isoelectric point (pI) values of the analyze proteins, ranging from 5.4 to 6.7 (Table 1), distinct differences were observed in the fluorescence responses for each protein. The results suggest that multivalent and differential interactions exist between PLL-Dnc and the analyze proteins through electrostatic interaction, hydrophobic interaction, and hydrogen bonding.

The resulting discriminative information was statistically evaluated by LDA, a statistical technique that converts input variables into a simple graphical output, where the inter- and intracluster variances are maximized and minimized, respectively.\(^{(26)}\) The LDA output of the training data matrix resulted in five canonical scores [Fig. 3(b)]. Among the obtained canonical scores, the first three scores were visualized in a three-dimensional plot, as shown in Fig. 3(c). The LDA score plot displays six well-separated clusters corresponding to the respective analytes with 100% accuracy according to the leave-one-out classification (also called jackknife

<table>
<thead>
<tr>
<th>Protein</th>
<th>Abbreviation</th>
<th>pI</th>
<th>Molecular weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin</td>
<td>ALB</td>
<td>5.7</td>
<td>67000</td>
</tr>
<tr>
<td>α1-Antitrypsin</td>
<td>AAT</td>
<td>5.4</td>
<td>44000</td>
</tr>
<tr>
<td>Fibrinogen</td>
<td>FIB</td>
<td>5.7</td>
<td>387000</td>
</tr>
<tr>
<td>Transferrin</td>
<td>TF</td>
<td>6.7</td>
<td>75000</td>
</tr>
<tr>
<td>α-Fetoprotein</td>
<td>AFP</td>
<td>5.5</td>
<td>69000</td>
</tr>
</tbody>
</table>
classification, which is a cross-validation where one point is left out for a test data matrix, while all other data are used to generate a training data matrix. Interestingly, ALB and AFP present large differences in their cluster positions in the LDA score plot despite their structural similarity, i.e., both comprise three domain structures with ~40% sequence homology.\(^{(27,28)}\) The high discrimination accuracy may have been achieved because the resulting fluorescence fingerprints reflect the complicated features of the PLL-Dnc-binding sites that are derived from the surface heterogeneity of the proteins (e.g., the type and steric arrangement of the amino acids at the binding site).

For further validation of the classification ability of the PLL-Dnc system, a blind test was performed. Additional 24 analytes [6 analytes (blank as well as 40 μg/mL ALB, AAT, FIB, TF, or AFP in CDCHO+ × 4 replicates) were prepared for a test data matrix, and the unknown analytes were classified on the basis of their Mahalanobis distances to the clusters generated from the training data matrix. Using this procedure, all unknown analytes were correctly identified with 100% accuracy, demonstrating the successful identification of hepatocyte-derived secretory proteins in a culture medium.
3.3 Sensing of ALB and AFP at various compositional ratios

To extend the potential of the array fabricated from PLL-Dnc, the system was applied to identify proteins with different concentrations and compositional ratios. A qualitative identification of analytes with different concentrations by LDA is usually used to evaluate the potential of a sensor system toward a quantitative analysis. Among the hepatocyte-derived secretory proteins, ALB and AFP were selected as analytes for the quantitative sensing, since these proteins are widely used as biomarkers in cultured hepatic cells and their progenitor cells. For example, in hepatocellular carcinoma-derived cell lines, ALB and AFP levels offer information on the activity of the cells and their growth cycle, which are altered upon treatment with therapeutic drugs. During the differentiation of pluripotent stem cells into hepatic cells, monitoring the ALB and AFP levels in the culture medium provides information on differentiation stages and the acquisition of a differentiated cellular function.

In this study, we prepared and analyzed ALB and AFP separately, and their mixtures in a culture medium at concentrations ranging from 0 to 2.0 μg/mL. The fingerprints of the fluorescence responses against the analytes were obtained and subsequently subjected to LDA. As visualized in Fig. 4(b), eight clusters corresponding to the respective analyte proteins were separated without overlap, giving a discrimination accuracy of 100% for the jackknife classification. In the blind test, additional 32 analytes (8 analytes (blank; 0.2, 1.0, and 2.0 μg/mL ALB in CDCHO+; 0.2, 1.0, and 2.0 μg/mL AFP in CDCHO+; mixture of 1.0 μg/mL ALB and AFP in CDCHO+) × 4 replicates) were prepared for a test dataset, and 30 out of 32 unknown analytes were correctly classified, providing an identification accuracy of 94%. Although both ALB and AFP clusters move along the positive direction on the score
(1) axis with increasing concentration, they move in the opposite direction along the score (2) axis. In addition, the LDA score plot shows that the cluster corresponding to a 1:1 mixture of ALB and AFP is located between the clusters of the respective proteins with the same total protein concentration. This result indicates that the PLL-Dnc system can recognize the different characteristics of these proteins and discriminate solutions containing different concentrations of proteins, which suggests that the PLL-Dnc system may potentially serve as a sensor to determine protein compositions. Further studies that concentrate on the use of, for example, support vector machine algorithms, which is a suitable analysis technique for datasets of numerical classes with nonlinear behavior, may potentially enable the PLL-Dnc system to quantify proteins in culture media.

Although the abundance of secretory biomarkers in culture media significantly depends on the conditions during cell culture (e.g., cell density, the surface area of the culture vessel, incubation time, and the volume of the culture medium), some previous studies have shown that hepatocytes cultured in vitro generally secrete ALB and AFP at concentrations ranging from 0.4 to 5.0 μg/mL in the cell culture supernatant. The concentrations of the stock sample solutions used in this study (≥2.0 μg/mL ALB and AFP in the culture medium) are comparable to the practical concentrations in culture supernatants, demonstrating the potential utility of the PLL-Dnc system for the analysis of biomarker proteins.

4. Conclusions

In conclusion, we have qualitatively detected a series of hepatocyte-derived secretory proteins spiked into a culture medium using an array made of environment-sensitive polymer PLL-Dnc. In addition to its high identification ability against secretory proteins in culture media, the sensor array based on PLL-Dnc presents many advantages, such as its facile synthesis, minimal sensor elements, and rapid output. The results of this work confirm the feasibility of this PLL-Dnc system for protein identification in real cell culture media that contain secretory proteins. The production of plasma proteins is a major function of hepatocytes, therefore, the identification of these proteins secreted into culture supernatants is useful for the quality control of cultured hepatocytes, i.e., for the evaluation of whether the cultured cells display the specialized function of the liver. Furthermore, monitoring secretory proteins in culture supernatants after the treatment of the hepatocytes with drugs suggests the potential of this system for use in the cell-based assessment of drug-induced hepatotoxicity. Thus, we believe that this PLL-Dnc system is a promising noninvasive and versatile method for the evaluation of cultured cells.

Acknowledgments

This work was supported by the Project Focused on Developing Key Evaluation Technology: Development of Platform Technology for Drug Discovery Applying Regenerative Medicine from the Japan Agency for Medical Research and Development, AMED, and JSPS KAKENHI grant 17H04884.
References