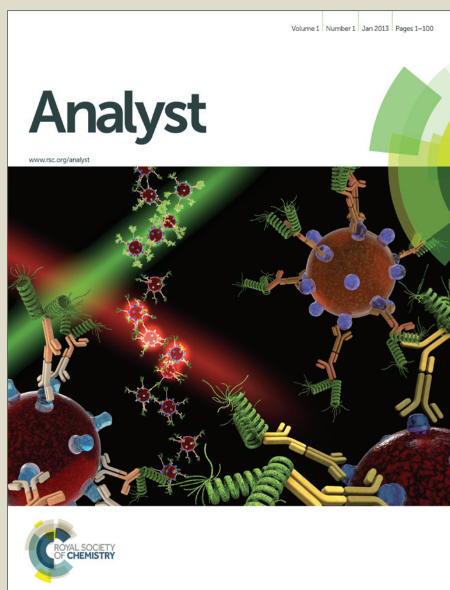
Analyst

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/analyst

Journal Name

RSCPublishing

COMMUNICATION

Enzymatic fingerprinting of structurally similar homologous proteins using polyion complex library constructed by tuning PEGylated polyamine functionalities

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Shunsuke Tomita,[*a] Tomohiro Soejima[b], Kentaro Shiraki[c] and Keitaro Yoshimoto[*a]

Human plasma proteins and even structurally similar homologous albumins were fingerprinted and discriminated by a sensor array consisting of a polyion complex library with artificial differentiation constructed by facile tuning of PEGylated polyamine functionalities.

Polyion complexes (PICs), *i.e.*, complexes between two oppositely charged polyelectrolytes formed *via* electrostatic interactions, have attracted a great deal of interest. This is partly due to the obviousness and simplicity of the general mechanism of PIC formation, which allows generation of materials with various desired properties. Well-known examples of PIC applications include membrane separation[1] and wound dressing technologies.[2] Charged amino acids are distributed on protein surfaces, and hence PICs can be formed between polyelectrolytes and counter-charged proteins, leading to new approaches in protein purification,[3] biosensor design,[4] and drug delivery.[5]

We found that enzymes were reversibly inhibited through PIC formation with polyelectrolytes.[6] By utilizing this reversible inhibition, an array-based approach to discriminate proteins by libraries of PICs between anionic enzymes and a cationic poly(ethylene glycol)-modified (PEGylated) polyamine was developed previously.[7] This approach is categorized as an "optical sensor array," where cross-reactive receptors in a sensor array interact differently with sensing targets and generate a specific response pattern that can be statistically analyzed to identify targets.[8] Over the past several years, optical sensor arrays for protein discrimination have been developed successfully.[9] Our previous approach provided high levels of structural diversity, a key feature to obtain differential patterns, for cross-reactive PICs with lower synthetic efforts by use of naturally occurring differentiation of enzymes, allowing identification of seven plasma protein solutions at 100 nM by only three PICs.[7]

To provide a more effective and practical system suitable for sensing of proteins and biochemical fluids with similar characteristics, a large set of cross-reactive PICs is required to select a combination of sensor elements capable of sufficient discrimination of the sensing targets. However, commercially available pairs of enzymes and substrates are limited, and therefore we set another goal, that of creating PIC libraries with artificial differentiation to extend the number of potential pairs of enzymes and PEGylated polyamines. The facile quaternization of PEGylated polyamine with functional halides attracted our attention because of its recent use for tuning the properties of polymeric materials, such as antibiotics,[10] drug carriers,[11] and drugs.[12] Following the previously described methodology,[11,13] two poly(ethylene glycol)-*block*-poly(*N*,*N*-dimethylaminoethyl methacrylate) (PEG-*b*-PAMA) with different molecular weights (**P1** and **P2**) and three quaternized PEG-*b*-PAMA (PEG-*b*-QPAMA) with different functional groups (**P3**, **P4**, and **P5**) were prepared in this study (Figure 1A, details are shown in Experimental section in ESI†).

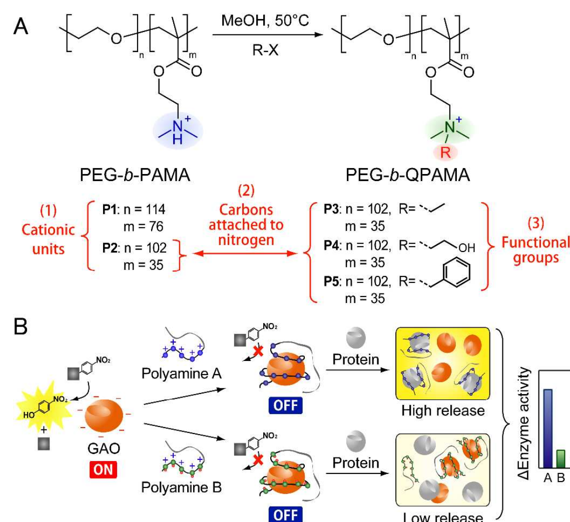


Fig. 1 (A) Strategy for the development of PEGylated polyamines with artificial differentiation. (B) Schematic illustration of decrease in GAO activity by PIC formation with PEGylated polyamines, and subsequent differential competitive interactions between PICs and a protein, generating a protein-specific pattern.

From the perspective of enzyme-based sensor arrays, anionic β-galactosidase from *Aspergillus oryzae* (GAO), an enzyme mainly used in the field of food chemistry,[14] is suited for the construction of a sensitive and accurate system as it has (i) high catalytic activity for hydrolysis of a conventional chromogenic substrate and (ii) high stability to withstand cryopreservation of its aqueous solution, enabling simplification of procedures. The sensing strategy is based on our recent findings,[6,7] where reversible PIC formation between GAO and five PEGylated polyamines was accompanied by a decrease in enzyme activity (Figure 1B). The subsequent mixing of analyte proteins with each PIC in an array causes enzyme release from PIC through competitive interactions, resulting in partial activity recovery. We postulated that different functionalities of PEGylated polyamines would provide diverse extents of enzyme recovery, generating protein-specific patterns that can be used for protein discrimination (Figure 1B).

Analyst Accepted Manuscript

To test our assumptions, titration experiments in 10 mM MOPS buffer (pH 7.0) were first performed to determine optimum GAO/PEGylated polyamine ratio (Figure 2). All polymers decreased GAO activity, but an interesting trend was found, where higher molecular weight (**P1** vs. **P2**) and quaternization (**P2** vs. **P3**, **P4**, and **P5**) provided a greater effect on the decrease in GAO activity. Considering the low pK_a of PAMA segment (~7.0, suggested from Figure S3, ESI†), the number of positive charges was increased nearly twofold by quaternization. Therefore, increased multiple electrostatic interactions of polyamine were responsible for the greater inhibitory effect. In addition, decreased activity of GAO did not follow the order of Log *P* values of R groups in PEG-*b*-QPAMAs, obtained by the program ALOPGs[15] (**P3**, 1.44; **P4**, –0.40; **P5**, 2.56), indicating that the hydrophobicity of introduced groups is not correlated with the inhibitory effect on GAO activity. Such different inhibitory effects of our PEGylated polyamines would provide diverse interactions between PICs and analyte proteins.

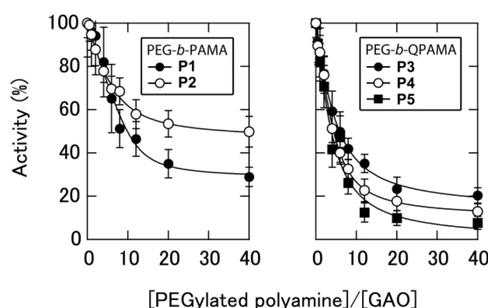


Fig. 2 Changes in activity of GAO. Titration of PEGylated polyamines to 1.0 nM GAO in 10 mM MOPS (pH 7.0).

From preliminary titration of proteins to prepared PIC (Figure S6, ESI†), the concentration of analyte proteins (5.0 μg/mL) was determined to evaluate the ability of PIC library possessing artificial differentiation of PEGylated polyamines. For the initial test, five the most abundant human plasma proteins were chosen as sensing targets—immunoglobulin G (IMM), fibrinogen (FIB), transferrin (TRA), albumin (HALB), and α_1-antitrypsin (ANT) (Figure S7, ESI†). Changes in the initial slope of Abs_{400} derived from enzymatic hydrolysis of substrates (v-v_0) for each PIC in a sensor array upon addition of each plasma protein at 5.0 μg/mL were recorded using a microplate reader, generating 150 data points (5 PICs × 5 proteins × 6 replicates) (Table S1, ESI†). These multivariate data were then analyzed by linear discriminant analysis (LDA), a statistical technique to simplify the data set obtained from a sensor array, allowing differentiation and classification of response patterns.[16] In LDA analysis, classification accuracy was initially calculated with the Jackknife classification procedure[17] to determine the minimal set of PICs for sufficient discrimination of plasma proteins. Although we observed accuracy of 60% – 83% using one PIC, 100% accuracy was achieved using a combination of only three PICs (GAO with **P2**, **P4**, and **P5**) (Table S2, ESI†). As shown in Figure 3A, the responses obtained by this PIC set roughly showed a negative correlation with experimentally determined p*I* of proteins, suggesting that the surface charges of proteins play an important role in protein–PIC interaction. Comparing FIB with TRA in the case of GAO/**P4**, however, the response of FIB was greater than that of TRA despite lower p*I* of TRA and lower Log *P* values of **P4**'s R groups compared to **P5**, indicating that responses were also influenced by various surface properties of proteins, such as heterogeneous distributions of polar and non-polar residues and morphological characteristics.

To visually evaluate the statistical significance of differences between patterns, discriminant scores of the first two discriminant functions obtained using LDA are plotted with ellipses representing confidence intervals (P = 0.68, ± 1 standard deviation (S.D.))[18] for the means of each plasma protein (Figure 3B, confidence intervals (P = 0.95, ± 2 standard error (S.E.)) are also shown in Figure S8, ESI†). All clusters were separated from each other, meaning accurate discrimination of analytes in accordance with the Jackknife classification. The detection efficiency was further validated by the identification of unknown samples according to their Mahalanobis distances to the respective groups with 100% accuracy (20 of 20) (Table S3, ESI†). Taken together, use of different cationic units, the number of carbons attached to nitrogen, and functional groups in PEGylated polyamines are effective to discriminate plasma proteins. Furthermore, interactions between PICs and proteins were translated and simultaneously amplified by catalytic reaction of enzymes, and hence, PICs with GAO were able to differentiate 5.0 μg/mL plasma proteins ranging from 13 nM for FIB (Mw: 387 kDa) to 113 nM for ANT (Mw: 44 kDa), comparable to the previously reported sensitive colorimetric array-based protein sensor.[9c]

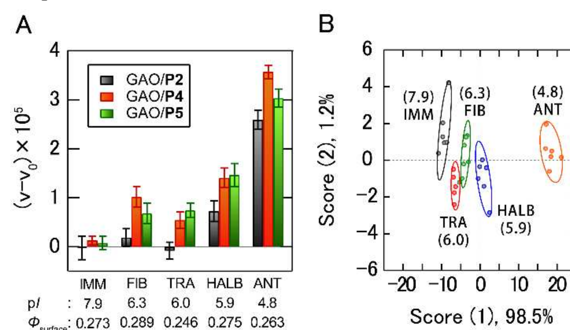


Fig. 3 Sensing of human plasma proteins using three PICs (GAO with **P2**, **P4**, and **P5**). (A) Enzyme activity patterns for five plasma proteins at 5.0 μg/mL. Each value represents the average of six parallel measurements with 1 S.D. p*I* obtained from the pH dependence of zeta-potential and the surface hydrophobicity ($\Phi_{surface}$) of proteins are shown (see Experimental section in ESI†). (B) Discriminant score plot of the first two discriminant functions of enzyme activity patterns analyzed by LDA. The ellipses represent confidence intervals (P = 0.68, ± 1 S.D.) for the individual plasma proteins. p*I* values of proteins are shown in parentheses.

After the successful discrimination of plasma proteins by the library of PICs between PEGylated polyamines and GAO, we investigated whether PIC sensor array could be used to discriminate structurally similar homologous proteins. Many albumins are allergenic, and antibodies are typically reactive to their homologs because of strong sequential and conformational similarity of albumins,[19] e.g., children with an allergy to cow's milk can develop sensitization to dog or cat epithelial albumins without any direct contact with the respective animals.[20] Thus, homologous albumins are challenging targets for our PIC sensor array in terms also of comparison with the immune system.

Therefore, we selected four albumins from different mammalian species—human (HALB), rabbit (RALB), equine (EALB), and bovine (BALB)—with sequence identity of about 70%,[19,21] and very close resemblances in tertiary structure,[21] molecular weight, and p*I* (Figure S7, ESI†). Albumins at a concentration of 5.0 μg/mL were analyzed using the same methodology as described for plasma proteins. Interestingly, the combination of PICs of GAO with **P1**, **P2**, and **P3** showed the best discrimination among all PICs with accuracy of 96% *via* Jackknife classification (Figure 4 and Table S4, ESI†), and 81% accuracy was observed in a blind test (13 of 16) (Table S5, ESI†). While the discriminant scores for albumins with very similar p*I*s (HALB, RALB, and EALB) were clustered closely, relatively acidic BALB was separated from the others (Figure 4B), which was similar to the case of plasma proteins (Figure 3). It should be noted that RALB and EALB were almost differentiated regardless of nearly identical p*I* and surface hydrophobicity. More detailed data are needed to elucidate the origin of these differences, but it confirmed that the array-based approach would be valuable for the discrimination of homologous proteins, as also indicated previously.[22]

Analyst Accepted Manuscript

 This journal is © The Royal Society of Chemistry 2012

Finally, to further evaluate the role of electrostatic interactions in pattern generation, first discriminant scores of eight proteins (plasma proteins and homologous albumins) analyzed by five PICs with LDA were compared with p*I*s (Figure S9). We observed that score appeared to correlate well with protein surface charges (correlation coefficient = –0.86), suggesting that responses obtained by our PIC library were mainly driven by electrostatics.

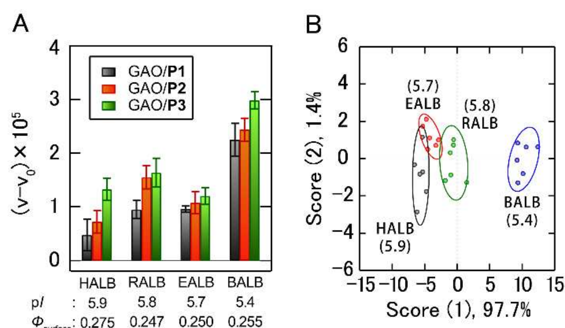


Fig. 4 Sensing of homologous albumins using three PICs (GAO with **P1**, **P2**, and **P3**). (A) Enzyme activity patterns for four homologous albumins at 5.0 μg/mL. Each value represents the average of six parallel measurements with 1 S.D. Experimentally determined p*I* and $\Phi_{surface}$ of proteins are shown. (B) Discriminant score plot of the first two discriminant functions of enzyme activity patterns analyzed by LDA. The ellipses represent confidence intervals ($P = 0.68$, ± 1 S.D.) for the individual homologous albumins. p*I* values of proteins are shown in parentheses.

In conclusion, a PIC library possessing artificial differentiation was constructed by tuning PEGylated polyamine functionalities, including cationic units, carbons attached to nitrogen, and functional groups. The selection of a minimal set of PICs from the library allowed the development of a simple yet effective array-based system for fingerprinting and discriminating human plasma proteins and structurally similar homologous albumins. Significantly, each PIC had less specificity for homologous albumins, but the combination of PICs achieved identification of homologous albumins that are sometimes misrecognized by the immune system,[19] indicating the power of the array-based approach. We expect that combining two strategies to construct cross-reactive PIC libraries—artificial differentiation of PEGylated polyamines and previously reported naturally occurring differentiation of enzymes[6]—will enable improving accuracy of homologous albumin discrimination, and moreover, increasing the number of proteins that can be differentiated by a single sensor array. From the viewpoint of analytical chemistry, enzyme-based sensor array is advantageous in terms of sensitivity because response patterns are determined from the rate of enzyme-catalyzed reaction, where background intensity can be neglected. Therefore, our method may be applicable to the analysis of biochemical samples such as human plasma or cell lysate by changing chromogenic substrate to fluorogenic one and/or increasing the number of cross-reactive PICs as necessary.

We thank Prof. Dr. Yukio Nagasaki and Shinpei Kudo (Graduate School of Pure and Applied Sciences, University of Tsukuba) for technical assistance in polymer synthesis and zeta-potential measurements. This work was supported by a Grant-in-Aid for JSPS Fellows, Young Scientists (B, 26810074), and Scientific Research (B, 24350037).

Notes and references

[a]Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo, 153-8902, Japan. Tel: +81-3-5454-6580; E-mail: s_tomita@bio.c.u-tokyo.ac.jp; ckeitaro@mail.ecc.u-tokyo.ac.jp
[b]College of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo, 153-8902, Japan.
[c]Faculty of Pure and Applied Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan.

† Electronic Supplementary Information (ESI) available: Experimentals, protein profiles, protein titration to PIC, response profiles, LDA analysis. See DOI: 10.1039/b000000x/

1 Q. Zhao, Q. F. An, Y. Ji, J. Qian and C. Gao, *J. Membrane. Sci.*, 2011, **379**, 19.
2 R. Jayakumar, M. Prabaharan, P. T. Sudheesh Kumar, S. V. Nair and H. Tamura, *Biotechnol. Adv.*, 2011, **29**, 322.
3 (a) P. McDonald, C. Victa, J. N. Carter-Franklin and R. Fahrner, *Biotechnol. Bioeng.*, 2009, **102**, 1141; (b) Y. Xu, M. Mazzawi, K. Chen, L. Sun and P. L. Dubin, *Biomacromolecules*, 2011, **12**, 1512.
4 (a) F. Lisdat, R. Dronov, H. Mohwald, F. W. Scheller and D. G. Kurth, *Chem. Commun.*, 2009, 274; (b) K. Ariga, Q. Ji, T. Mori, M. Naito, Y. Yamauchi, H. Abe and J. P. Hill, *Chem. Soc. Rev.*, 2013, **42**, 6322.
5 (a) Y. Lee, T. Ishii, H. J. Kim, N. Nishiyama, Y. Hayakawa, K. Itaka and K. Kataoka, *Angew. Chem. Int. Ed.*, 2010, **49**, 2552; (b) D. S. Manickam, A. M. Brynskikh, J. L. Kopanic, P. L. Sorgen, N. L. Klyachko, E. V. Batrakova, T. K. Bronich and A. V. Kabanov, *J. Control. Release*, 2012, **162**, 636.
6 (a) S. Ganguli, K. Yoshimoto, S. Tomita, H. Sakuma, T. Matsuoka, K. Shiraki and Y. Nagasaki, *J. Am. Chem. Soc.*, 2009, **131**, 6549; (b) S. Tomita, L. Ito, H. Yamaguchi, G. Konishi, Y. Nagasaki and K. Shiraki, *Soft Matter*, 2010, **6**, 5320; (c) S. Tomita and K. Shiraki, *J. Polym. Sci. Part A: Polym. Chem.*, 2011, **49**, 3835; (d) T. Kurinomaru, S. Tomita, S. Kudo, S. Ganguli, Y. Nagasaki and K. Shiraki, *Langmuir*, 2012, **28**, 4334.
7 S. Tomita and K. Yoshimoto, *Chem. Commun.*, 2013, **49**, 10430.
8 J. R. Askim, M. Mahmoudi and K. S. Suslick, *Chem. Soc. Rev.*, 2013, **42**, 8649.
9 (a) M. De, S. Rana, H. Akpinar, O. R. Miranda, R. R. Arvizo, U. H. F. Bunz and V. M. Rotello, *Nat. Chem.*, 2009, **1**, 461; (b) O. R. Miranda, H. T. Chen, C. C. You, D. E. Mortenson, X. C. Yang, U. H. Bunz and V. M. Rotello, *J. Am. Chem. Soc.*, 2010, **132**, 5285; (c) X. Li, F. Wen, B. Creran, Y. Jeong, X. Zhang and V. M. Rotello, *Small*, 2012, **8**, 3589; (d) Y. Tao, X. Ran, J. Ren, and X. Qu, *Small*, 2014, DOI: 10.1002/smll.20140061.
10 (a) S. Venkataraman, Y. Zhang, L. Liu and Y. Y. Yang, *Biomaterials*, 2010, **31**, 1751; (b) T. R. Stratton, B. M. Applegate and J. P. Youngblood, *Biomacromolecules*, 2011, **12**, 50.
11 (a) A. Tamura, M. Oishi and Y. Nagasaki, *J. Control. Release*, 2010, **146**, 378; (b) M. Tamura, S. Ichinohe, A. Tamura, Y. Ikeda and Y. Nagasaki, *Acta. Biomater.*, 2011, **7**, 3354.
12 Y. Nagasaki, T. Yamazaki, A. Kikuchi and M. Harada-Shiba, *Colloids. Surf. B. Biointerfaces*, 2014, **113**, 237.
13 T. Ishii, H. Otsuka, K. Kataoka, and N. Nagasaki, *Langmuir*, 2004, **20**, 561.
14 (a) N. Albayrak and S. T. Yang, *Biotechnol. Bioeng.*, 2002, **77**, 8-19.; (b) R. Gaur, H. Pant, R. Jain and S. K. Khare, *Food Chemistry*, 2006, **97**, 426-430.
15 I. V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V. A. Palyulin, E. V. Radchenko, N. S. Zefirov, A. S. Makarenko, V. Y. Tanchuk and V. V. Prokopenko, *J. Comput. Aid. Mol. Des.*, 2005, **19**, 453.
16 P. Jr Anzenbacher, P. Lubal, P. Bucek, M. A. Palacios and M. E. Kozelkova, *Chem. Soc. Rev.*, 2010, **39**, 3954.
17 B. G. Tabachnick and L. S. Fidell, Using Multivariate Statistics, Allyn& Bacon, Needham Heights, MA, 4th edn, 2001.
18 C. David, A. Schleicher, W. Zuschratter and J. F. Staiger, *Eur. J. Neurosci.*, 2007, **25**, 2329.
19 M. Chruszcz, K. Mikolajczak, N. Mank, K. A. Majorek, P. J. Porebski and W. Minor, *Biochim. Biophys. Acta.*, 2013, **1830**, 5375.

Analyst Accepted Manuscript

20 J. Vicente-Serrano, M. L. Caballero, R. Rodríguez-Pérez, P. Carretero, R. Pérez, J. G. Blanco, S. Juste and I. Moneo, *Pediatr. Allergy Immunol.*, 2007, **18**, 503.
21 K. A. Majorek, P. J. Porebski, A. Dayal, M. D. Zimmerman, K. Jablonska, A. J. Stewart, M. Chruszcz and W. Minor, *Mol. Immunol.*, 2012, **52**, 174.
22 H. Kong, H. Wang, S. Zhang and X. Zhang, *Analyst*, 2011, **136**, 3643.

Analyst Accepted Manuscript

This journal is © The Royal Society of Chemistry 2012