Paper

Analyzing TOF-SIMS spectra of biopolymer using multivariate curve resolution

Satoka Aoyagi,[1,2*] Masayuki Okamoto,[2,3] Nobuhiko Kato[4] and Masahiro Kudo[4]
[1]*Faculty of Life and Environmental Science, Shimane University, 1060 Nishikawatsu-cho, Matsue, Shimane, 690-8504, Japan*
[2]*Division of Production Environment Engineering, United Graduate School of Agricultural Sciences, Tottori University, 4-101 Koyama-Minami, Tottori 680-8553, Japan*
[3] *Kao Corporation, 1334 Minato, Wakayama-shi, Wakayama 640-8580, Japan*
[4] *Faculty of Science and Technology, Seikei University, 3-3-1, Kitamachi, Kichijioji, Musashino, Tokyo, 180-8633, Japan*
* aoyagi@life.shimane-u.ac.jp

(Received : October 3, 2010; Accepted : January 7, 2011)

Fragment ion analysis of protein samples is crucial for structural evaluation of proteins. Interpretation of fragment ions from bioplolymers, however, is often difficult because fragmentation mechanisms in secondary ion mass spectrometry have not been clarified. Therefore employment of chemometrics is necessary to obtain detailed information from time-of-flight secondary ion mass spectrometry (TOF-SIMS) data of complex samples. In this study, multivariate curve resolution (MCR) was applied to extract a pure component spectrum from a TOF-SIMS spectrum which contains mixed materials information. The model biopolymer, chicken egg white lysozyme was immobilized on a gulutaraldehyde-activated aminosilanized indium-tin oxide (ITO) glass plate by covalent bonding. A reference sample such as substrates without the protein were prepared. All samples were measured with TOF-SIMS using Bi_3^+, and then the TOF-SIMS spectra data were analyzed using MCR with and without data preprocessing. As a result, images and spectra related to each material in the sample were obtained by MCR. The fragment ions in the extracted spectra are consistence with chemical information of each material.

1. Introduction

Interpretation of time-of-flight secondary ion mass spectrometry (TOF-SIMS) spectra is crucial for evaluating complex samples of macromolecules such as protein [1,2] and polymer materials containing complicated additives. Peak overlapping is one of the most difficult issues to interpret complicated TOF-SIMS spectra, especially for analyzing complex polymer samples such as proteins. Many protein-related peaks are related to other organic materials even contaminants.

Multivariate analysis (MVA) techniques have been employed to interpretation of TOF-SIMS spectra [3]. The most popular MVA technique is principal component analysis (PCA) which often provides useful information for characterizing TOF-SIMS spectra of complex samples. However, principal components are obtained by mathematical processes regarding eigenvalues and eigenvectors of variance-covariance matrices which require rotation of the raw data, and therefore PCA does not always provide meaningful information in terms of physics or chemistry [4].

On the other hand, multivariate curve resolution (MCR) [4-6] provides pure components spectra when it performs appropriately. Since clarifying fragment ions specific to a particular material is essential in evaluation of macromolecules, a pure component spectrum extracted by MCR is useful for selecting important fragment ions of each material in a sample. When spectra of pure components can be extracted from raw data, the analysis technique will be useful for evaluation of protein complex samples for studying protein-protein interaction.

In this study, a protein-immobilized sample was prepared as a model sample for MCR application.

2. Materials and Methods

Protein Sample preparation

An indium tin oxide (ITO) coated glass slide

(SIGMA- Aldrich Co., St Louis, MO) of 8 x 8 mm was aminosilanized with aminopropyltrimethoxysilan (Tokyo Kasei, Tokyo, Japan). The aminosilanized ITO glass plates were activated by glutaraldehyde and then soaked in a 0.1M-phosphate buffered saline (PBS) solution at pH 7.4 containing egg white lysozyme (Sigma-Aldrich) and allowed to react in the dark for 30 hr at 277 K, and then the protein molecules were immobilized on the ITO plates at amino groups by covalent bonding [7, 8]. After the ITO glass plates were washed in the PBS solution and then washed in pure water, sonic waves were applied for 10 s to remove adsorbed protein molecules. The glutaraldehyde-activated aminosilanized ITO glass plates without the protein was prepared as a reference sample. These samples were dried with a freeze dryer (VD-250F, Taitech, Saitama, Japan) before TOF-SIMS measurement.

TOF-SIMS measurement

Positive ion spectra of the protein sample and the reference sample were obtained with TOF-SIMS 5 (ION-TOF GmbH, Münster) with 25 keV Bi_3^{++} primary ion sources. All measurements were acquired while maintaining the primary ion dose at less than 10^{12} ions/cm2 to ensure static conditions.

MCR

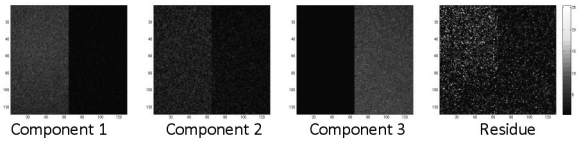
Peaks of secondary ions of the protein samples obtained with TOF-SIMS 5 were auto searched with the TOF-SIMS analysis software provided by ION-TOF GmbH, and then 308 peaks, ranging from m/z 12 to 326, were selected. Secondary ion images of the selected peaks were converted into binary image files (BIF) and they were transformed into matrix data using MIA_Toolbox (Eigenvector Research Inc., WA).

The matrix data were analyzed using the MCR program developed by S. Muto et al. [4]. The program is coded on Matlab (The MathWorks Inc., MA). The modified alternating least square (MALS) [5] was adopted to the program. The TOF-SIMS data of the protein sample were treated with Poisson-scaling (root mean scaling) [9-12] before the MCR calculations.

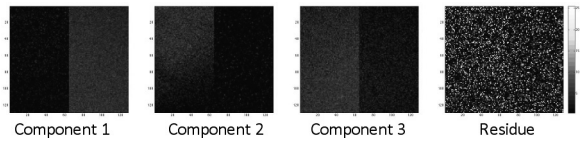
3. Results and discussion

The secondary ion figures of the lysozyme-immobilized sample and a reference sample without lysozyme were integrated into one figure as shown in Fig.1, and then the figure was analyzed with MCR. The left half of each figure in Fig. 1 represents the lysozyme-immobilized data and the right half one represents the reference sample without the protein. When an appropriate spectrum specific to the protein is obtained with MCR, the total secondary ion image of the spectrum appears only on the left half of the figure. Thus MCR results adequacy can be indicated by images.

Several combinations of a lysozyme sample datum and a reference sample datum have been examined in order to ensure similar results obtained by MCR with and without data preprocessing, Poisson-scaling. When the number of components is assumed to be three, the lysozyme sample area was divided into two components and the reference sample area was separated from them in both cases with and without Poisson-scaling. The main difference between with and without the data preprocessing is shown in the residues images. The residue image with Poisson-scaling may represent background or errors because it is homogeneous. On the other hand, the residue without data preprocessing shows clearly higher intensities in the lysozyme sample area than in the reference area, which indicates information related to the lysozyme sample still remains.



(a) Without data preprocessing

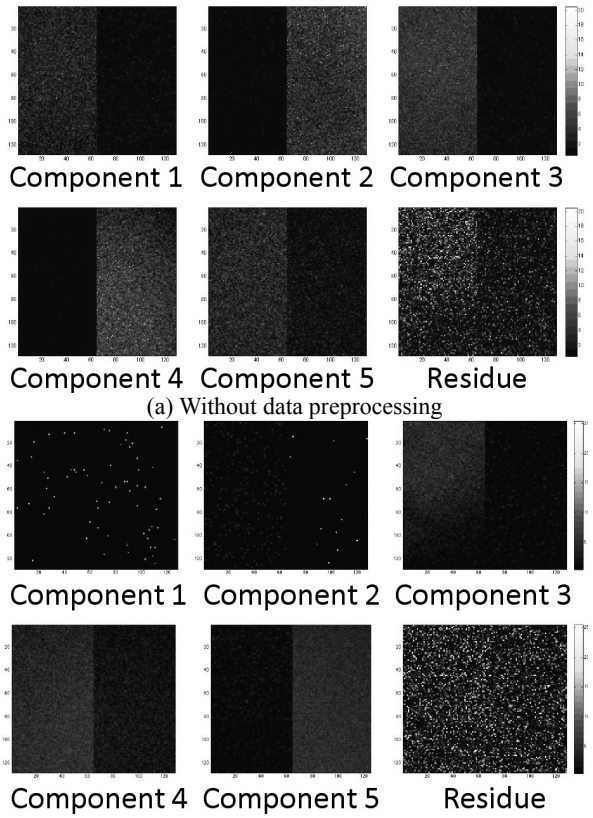


(b) With Poisson-scaling

Fig. 1. Secondary ion images of the components extracted by MCR when the number of components is assumed to be three. The left half of each figure is from TOF-SIMS data of the lysozyme-immobilized sample and the other half is from those of the reference sample (glutaraldehyde-activated aminosilanized ITO-coated glass substrate). Field of view is 120 μm.

MCR was performed on the TOF-SIMS data with assuming varying the number of components, which should be initially given in MCR calculation, from three to five. When the number is more than three, clear differences between results, with and without data preprocessing, appear as shown in Fig. 2. Regardless of the number of the components from three to five, the same number of components was extracted from

each area with Poisson-scaling. The number of components extracted from each area, however, increased when data preprocessing was not employed. In addition, residues with Poisson-scaling are homogeneous though residues without data preprocessing are inhomogeneous, which may suggest that the residue without data preprocessing still remains information on the sample and that the residue with Poisson-scaling represents mainly errors and background signals.



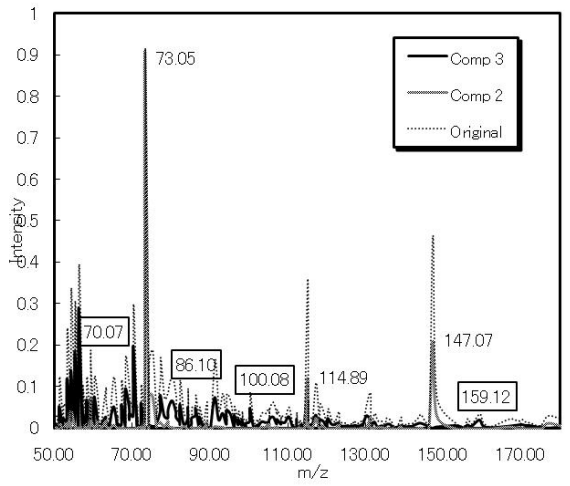
(a) Without data preprocessing

(b) With Poisson-scaling

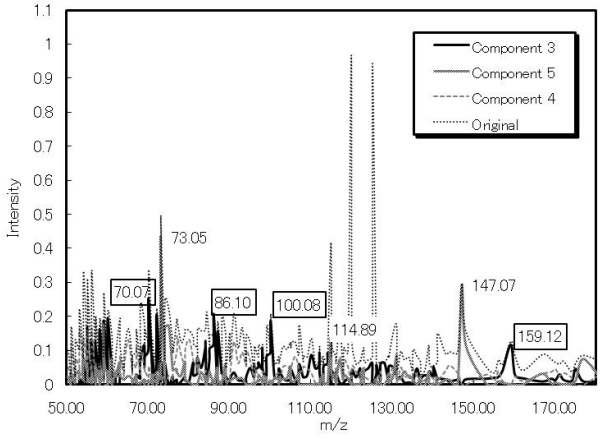
Fig. 2. Secondary ion images of the components extracted by MCR when the number of components is assumed to be five. The left half of each figure is from TOF-SIMS data of the lysozyme-immobilized sample and the other half is from those of the reference sample (ITO-coated glass substrate). Field of view is 120 μm.

A spectrum of each component with or without data preprocessing was evaluated whether it contained secondary ions specific to the material suggested by its distribution. It was found the spectrum of each component is consistent with a main material indicated by its image independent on the number of components initially given for MCR with or without data preprocessing. For instance, the spectra of component 2 in Fig. 3(a) and component 5 in Fig. 3 (b) show *m/z* 73.05 and 147.07 peaks related to glass and *m/z* 114.89 from In, and their intensities on the spectra of the components related to the lysozyme sample were almost zero.

On the other hand, both the spectra of component 3 in Fig. 3 (a) and (b) contain most of the protein-related amino acid fragments [14,15]. However, the lysozyme-related spectrum indicated by MCR with Poisson-scaling (the component 3 in Fig. 3(b)) contains more protein-related fragment ions than without data preprocessing. The detected peaks from the protein fragments in Table 1 are *m/z* 30.03, 43.03, 44.05, 60.04, 68.05, 69.03, 70.03, 70.07, 72.08, 73.06, 74.06, 81.04, 82.05, 84.04, 84.08, 86.10, 87.06, 88.04, 100.09, 102.06, 107.05, 110.07, 112.09, 120.08, 127.10, 130.07, 136.08, 159.09, and 170.06, and their intensities on the spectrum related to the reference are almost zero.



(a) Without data preprocessing



(b) With Poisson-scaling

Fig. 3. TOF-SIMS spectra of components related to the lysozyme-immobilized sample and the reference sample, respectively, extracted by MCR. Component 3 in (a) and (b) is related to lysozyme and components 2 in (a) and 5 in (b) are related to the reference sample.

Since the other components have less information in terms of indicating a particular material because they contain weak peaks, they

are omitted. The component 4 in Fig. 3 (b) shows less intensity in terms of the famous protein-related peaks [14,15] and its image in Fig. 2 (b) shows distribution not only on the left half but also on the right side which is the control sample area. Therefore, it is indicated that this component may be related to the organic part of the glutaraldehyde-activated aminosilanized ITO glass surface. Fig. 4 shows original secondary ion images of the protein-related peaks, and shows similar distributions to the image of component 3 in Fig. 2 (b). In addition, the intensity gradation of the component 3 image in Fig. 2 (b) and Fig. 4 may be depending on inhomogeneous immobilization of the protein which sometimes occurs partially.

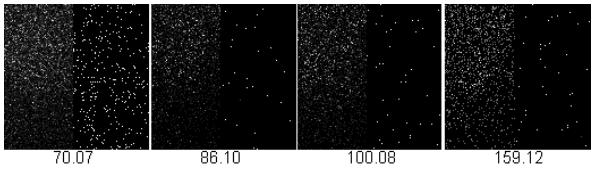


Fig. 4. Secondary ion images of the protein-related peaks, m/z 70, 86, 100 and 159. The left half of each figure shows the lysozyme-immobilized sample and the right one shows the control sample.

Table 1. Reported fragment ions related to the amino acids. [14,15]

Residues	Formula	m/z	Residues	Formula	m/z
Gly	CH4N	30.034	Lys	C5H10N	84.08
Arg	CH3N2	43.03	Ile, Leu	C5H12N	86.10
Ala	C2H6N	44.05	Asn	C3H7N2O	87.06
Cys	CHS	44.98	Asn, Asp	C3H6NO2	88.04
Ser	C2H6NO	60.045	Asn	C4H4NO2	98.02
Met	C2H5S	61.011	Arg	C4H10N3	100.09
Pro	C4H6N	68.05	Arg	C4H11N3	101.10
Thr	C4H5O	69.034	Glu	C4H8NO2	102.06
Asn	C3H4NO	70.029	Tryr	C7H7O	107.05
Pro	C4H8N	70.066	His	C5H8N3	110.07
Ser	C3H3O2	71.013	Arg	C5H10N3	112.09
Val	C4H10N	72.081	Phe	C8H10N	120.08
Arg	C2H7N3	73.064	Arg	C5H11N4	127.10
Thr	C3H8NO	74.061	Trp	C9H8N	130.07
His	C4H5N2	81.045	Phe	C9H7O	131.05
His	C4H6N2	82.053	Tryr	C8H10NO	136.08
Val	C5H7O	83.05	Trp	C10H11N2	159.09
Gln, Glu	C4H6NO	84.045	Trp	C11H8NO	170.06

Thus it is indicated that a spectrum extracted by MCR is useful to find fragment ions from a particular macromolecule, which is often difficult by manual analysis. Further study is necessary to establish MCR application protocols for TOF-SIMS analysis. This technique can be applied to evaluation of complicated samples containing complex molecules such as biochips and polymer materials.

4. Conclusions

MCR results indicated that TOF-SIMS spectra of the protein sample are divided into a spectrum of each material in the sample. Extraction of a pure component spectrum from TOF-SIMS data is useful to find important peaks of fragment ions specific to a particular molecule, which strongly supports TOF-SIMS spectra interpretation of complex samples.

Acknowledgement

The authors thank Professor Shunsuke Muto, Nagoya University, for providing his program of MCR.

References

[1] S. Aoyagi, *Surf. Inter. Anal.* **41**, 136 (2009).
[2] S. Aoyagi, M. Higuchi, N. Kato, and M. Kudo, *e-J. Surf. Sci. Nanotech.* 7**,** 715(2009).
[3] J.L.S. Lee, B.J. Tyler, M.S. Wagner, I.S. Gilmore and M.P. Seah, *Surf. Inter. Anal.* **41**, (2009) 76.
[4] S. Muto et al., *Mater. Trans*., **50** (2009) 964.
[5] J.-H. Wang et al., *Anal. Chim. Acta* **476** (2003) 93.
[6] J. Jaumot, R. Gargallo, A. de Juan, and R. Tauler, *Chemometr. Intell. Lab. Syst*. **76** (2005) 101.
[7] S. Aoyagi, R. Imai, K. Sakai and M. Kudo, *Biosens. Bioelectron.*, **18** (2003) 791
[8] Hopwood, *Histochem. J*. **4** (1972) 267
[9] B.J. Tyler, G. Rayal, D.G. Caster, *Biomaterials*, **28** (2007) 2412.
[10] J.L.S. Lee, I.S. Gilmore, I.W. Fletcher and M.P. Seah, *Surf. Interface Anal*., **41** (2009) 653.
[11] M.R. Keenan and P.G. Kotula, *Surf. Interface Anal*., **36** (2004) 203.
[12] M.R. Keenan and P.G. Kotula, *Appl. Surf. Anal*., **231-232** (2004) 240.
[14] J-B. Lhoest, E. Detrait, PvdBd Aguilar and P. Bertrand: J. Biomed. Materials Res. **41**, 95 (1998)
[15] D. S. Mantus, B. D. Ratner, B. A. Carlson, and J. F. Moulder, Anal. Chem. **65**, 1431 (1993)