Multivariate ToF-SIMS image analysis of polymer microarrays and protein adsorption

Andrew L. Hook, Philip M. Williams, Morgan R. Alexander, and David J. Scurr

Articles you may be interested in

Latest applications of 3D ToF-SIMS bio-imaging
Biointerphases **10**, 018902 (2015); 10.1116/1.4907727

Study of the interactions of proteins with a solid surface using complementary acoustic and optical techniques
Biointerphases **9**, 029015 (2014); 10.1116/1.4874736

An equipment-free polydimethylsiloxane microfluidic spotter for fabrication of microarrays
Biomicrofluidics **8**, 026501 (2014); 10.1063/1.4871935

A coarse grain model for protein-surface interactions
J. Chem. Phys. **139**, 095102 (2013); 10.1063/1.4819131

Self-assembled monolayers for polymer and protein cationization with time-of-flight secondary ion mass spectrometry
J. Vac. Sci. Technol. A **18**, 1114 (2000); 10.1116/1.582309

Multivariate ToF-SIMS image analysis of polymer microarrays and protein adsorption

Andrew L. Hook, Philip M. Williams, Morgan R. Alexander, and David J. Scurr[a)]

Laboratory of Biophysics and Surface Analysis, School of Pharmacy, University of Nottingham, Nottingham NG7 2RD, United Kingdom

The complexity of hyperspectral time of flight secondary ion mass spectrometry (ToF-SIMS) datasets makes their subsequent analysis and interpretation challenging, and is often an impasse to the identification of trends and differences within large sample-sets. The application of multivariate data analysis has become a routine method to successfully deconvolute and analyze objectively these datasets. The advent of high-resolution large area ToF-SIMS imaging capability has enlarged further the data handling challenges. In this work, a modified multivariate curve resolution image analysis of a polymer microarray containing 70 different poly(meth)acrylate type spots (over a 9.2×9.2 mm area) is presented. This analysis distinguished key differences within the polymer library such as the differentiation between acrylate and methacrylate polymers and variance specific to side groups. Partial least squares (PLS) regression analysis was performed to identify correlations between the ToF-SIMS surface chemistry and the protein adsorption. PLS analysis identified a number of chemical moieties correlating with high or low protein adsorption, including ions derived from the polymer backbone and polyethylene glycol side-groups. The retrospective validation of the findings from the PLS analysis was also performed using the secondary ion images for those ions found to significantly contribute to high or low protein adsorption. © 2015 Author(s). *All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License.* [http://dx.doi.org/10.1116/1.4906484]

I. INTRODUCTION

The application of time of flight secondary ion mass spectrometry (ToF-SIMS) analysis has been successfully applied throughout a range of scientific disciplines for the analysis of both inorganic and organic sample types.[1–3,6] For chemical analysis by ToF-SIMS, hyperspectral images are produced where each individual pixel contains a full mass spectrum. A typical image rastered over 500 μm $\times$ 500 μm at 256×256 resolution results in 65 536 separate spectra, which themselves contain hundreds of ion peaks[4] of significant intensity. For example, if the spectrum were to contain 1000 ions of interest this would lead to $>65 \times 10^6$ separate data points for a typical sampled area. Automated multiple area acquisition and image stitching now routinely produces images of tens of millimeters, which proportionally increases the dataset size. Traditionally, the complexity of these datasets has proven a barrier to the uptake of the technique; however, the increasing application of multivariate analysis has significantly improved the ability to interpret successfully these datasets as detailed in a recent review by Graham *et al.*[4] and previously applied to ToF-SIMS and other mass spectrometry data.[5,6]

One sample type particularly relevant to ToF-SIMS chemical imaging is the polymer microarray system, whereupon hundreds to thousands of unique polymers are printed onto a substrate at addressable locations.[7–10] Polymer microarrays have become a key tool for materials discovery projects and have been widely used to assess a number of material–biological applications including stem cells,[11–13] bacteria,[14–16] and cell sorting.[17–20] The large number of material interactions that can be rapidly assessed using polymer microarrays has resulted in surface structure–property models that describe the biological performance of materials based upon a measured physical property. To achieve this, high throughput surface characterization of polymer microarrays has been a necessary development and has been achieved for various techniques including x-ray photoelectron spectroscopy, ToF-SIMS[8,9] water contact angle (WCA)[21,22] atomic force microscopy,[23] surface plasmon resonance,[24] and force measurements.[25] Partial least square (PLS) regression has been highly successful at correlating a univariate property, such as cell number, with a multivariate data set, such as the hundreds of secondary ions produced in ToF-SIMS spectra.[26,27] This method was initially validated for a large polymer library by linking ToF-SIMS spectra with WCA,[28] and has been successfully applied to predict the stem cell attachment to polymers[11,12] and bacterial attachment[14] from the chemical information represented in ToF-SIMS spectra. Importantly, these have identified relationships between surface chemistry and various biological responses. As technical advances in ToF-SIMS enable the analysis of increasingly complex organic samples, including polymers[29] and biological materials,[30] the multivariate data analysis (MVA) techniques used to assess the data must also be advanced. Large area imaging has increased the potential size of the hyperspectral imaging datasets, for example, a relatively modest analysis area of 10×10 mm imaged at a resolution of 500 pixels

a)Electronic mail: david.scurr@nottingham.ac.uk

per mm would comprise 25×10^6 pixels. With the example of 1000 secondary ions of interest this would result in a total number of data-points of 25×10^9. In this range, the ability to perform MVA analysis is hindered by conventional computing systems and data processing methodologies.

The potential of employing high-performance computing to undertake multivariate curve resolution (MCR) image analysis has been previously demonstrated by the authors,[31] as applied to ToF-SIMS data of eight polymer microarray printed spots combined to form a single composite hyperspectral image (consisting of a 524 288 pixel area). In this present study, a modified MCR imaging analysis approach has allowed the successful deconvolution of a polymer microarray consisting of 70 distinct polymer chemistries without a need for postacquisition data binning or a reduction in the acquired spatial resolution. Importantly, for the wider uptake of this method, this modified approach can be performed using a high-specification (8–16 GB, multicore) desktop computer in addition to a high performance compute cluster.

II. EXPERIMENT

In this paper, a poly(meth)acrylate microarray consisting of 70 unique polymer spots was contact printed with replicates and was subsequently analyzed using ToF-SIMS and a protein binding assay. The ToF-SIMS data for the native array were analyzed using MCR image analysis. The ToF-SIMS and protein adsorption data for each individual polymer spot were then correlated using PLS regression analysis.

A. Microarray printing

Polymer microarrays were formed as previously described in detail elsewhere[32,33] using a XYZ3200 dispensing workstation (Biodot). Epoxy-functionalized glass slides (Genetix) were dip coated in 4% (w/v) poly(hydroxy ethylmethacrylate) (pHEMA) (Sigma, cell culture tested) in ethanol at a withdrawal rate of approximately 30 mm/s. The slides were held horizontally for 1 min to allow solvent evaporation and then placed in a drying rack for 3 days. Polymerization solution composed of 75% (v/v) monomer (Sigma) in dimethylformamide (DMF) with 1% (w/v) photoinitiator 2,2-dimethoxy-2-phenylacetophenone was printed onto the pHEMA coated slides using 946PM6B pins (ArrayIt) at $O_2 < 2000$ ppm, 25 °C, 40% humidity. After printing each material, slides were irradiated with a long wave UV source for 30 s. Slides were irradiated for a further 10 min once all materials had been printed. Once array formation was complete the slides were vacuum extracted at <50 mTorr for 7 days. The specific chemistry and relevant CAS number for the 71 spots printed is shown in the supplementary Table SI.1.[34]

B. Protein adsorption

Polymer microarrays were immersed in 25 μg/ml tetramethylrhodamine isothiocyanate labeled albumin (Sigma) in phosphate buffer saline (Gibco) at pH 7.4. The array was incubated with protein for 1 h under stagnate conditions at 37 °C before washing twice for 1 min with ultrapure water (18.2 MΩ), blotting, and overnight drying. Fluorescence images of the slide were acquired using an array scanner (Genepix) using a 532 nm laser, 5 μm pixel size. The slide was also imaged prior to protein adsorption to measure background fluorescence. The fluorescence from each polymer spot was quantified by subtracting the fluorescence from an array going through the same procedure but with no protein from the fluorescence measured after protein adsorption from eight replicate samples.

C. ToF-SIMS analysis

ToF-SIMS analysis of the native array was performed using an IONTOF (GmbH) ToF-SIMS IV instrument utilizing a 25 keV Bi_3^+ primary ion source. An area of 9.2×9.2 mm was analyzed using the macroraster (large area) scanning facility encompassing the entire 70 polymer spot microarray in the "high-current bunched" mode. Data were acquired with a single scan of the analysis area at a resolution of 100 pixels per mm and 15 pulses per pixel. Owing to the insulating nature of the sample, charge compensation was applied in the form of a low energy ($\sim$20 eV) electron flood gun. Both positive and negative secondary ion data was collected; however, for brevity, only the negative data will be discussed in the work.

ToF-SIMS data analysis was performed using SurfaceLab6 (IONTOF GmbH) instrument software. In order to identify secondary ion peaks predominantly relating to the polymer chemistry, the 70 spot areas were selected using the region of interest (ROI) polyline selection tool and the total ion image data in the SurfaceLab6 instrument software. This data was used to generate a peak list comprising all secondary ions of significant intensity. The peak list was generated using the search function in SurfaceLab6 instrument software with a selection threshold of >100 counts. This peak list of 706 ions was then used to retrospectively rebuild the secondary ion images and to export specific ion intensities for each independent spot region. The secondary ion images were then Poisson corrected (SurfaceLab6 instrument software) and exported as text files for MCR analysis. However, where possible an advanced deadtime correction should be applied to such data.[35] The ion intensities for the individual polymer spots were exported and normalized to total ion count values prior to subsequent PLS analysis.

III. MULTIVARIATE DATA ANALYSIS

A. Multivariate curve resolution image analysis

MCR analysis was performed using a modified version of the Alternating Least Squares MCR package (MCR-ALS version 0.0.4)[36] in the R software environment (R version 3.1.0)[37] on a Dell C6220 computer server (a head node of the University of Nottingham's High-Performance Computer), with two 8-core processors (Intel Sandybridge E5–2670 2.6 GHz) and 128 GB RAM running Scientific Linux release 6. This single node could be considered as

equivalent to the processing power of a high-specification (large memory, multicore) desktop computer. The Poisson-corrected[38] ToF-SIMS image data of 706 ion intensities from the array of 920×920 positions amounted to 4.45 GB of 64-bit data. Whilst R and the MCR-ALS package can process such data on this single node, it requires up to 60 GB of memory to do so and runs unacceptably slowly, with each of often 50 or so cycles of alternating MCR taking over 10 min to complete. To overcome these restrictions we implemented the workflow below to partition the analysis across smaller, independent compute tasks, which can be computed in parallel on a compute cluster where available, or serially on a single computer as done in this example.

The alternating least-squares approach to MCR iterates the solution of "scores" (abundance or class maps) then "loadings" (ion intensities) until the residual sum-of-the-square (RSS) difference between experimental and model data fails to fall significantly further. The first iteration requires a guess of either the component scores maps or their ion loadings, and the success of MCR-ALS to derive the separate component scores and loadings depends on the quality of this initial estimate. Extracting a large number of components from such data often leads to computational errors (numerical singularities) unless the initial estimate is good. To extract the components from the data, we first recognized that the location of the components in the map was not initially a concern, so we generated a sampled subset of the data as 115×115 spectra of 706 ions by summing each 8×8 pixel block in the original data. This combined data contains all of the total ion counts and hence all of the components of the original (the spatial distribution within these combined areas is recovered later). We first performed MCR-ALS on this reduced pixel data (1/64th the size of the original) to find two components, which was readily achieved with random starting scores and loadings in less than 60 s, this process is illustrated in the supplementary material as Fig. SI-1(a). The two maps were then segmented to identify the location of the components by maximum-entropy thresholding to generate binary image maps of where each component was present in the scan [Fig. SI-1(b)]. We assumed that the map that contained the most number of separate domains of at least 50% the size of the largest in their map was that most likely to reflect the presence of at least two components. The map was split into two generating a third [Fig. SI-1(c)]. These three maps were then used for a new round of MCR-ALS [Fig. SI-1(d)], this time searching for three components. The complete cycle could be repeated over 100 times with each application deconvolving the data into one more component. As each application for MCR-ALS is starting from a reasonable estimation of the final scores distribution, the time taken to complete each round was also in the order of 60 s.

Having determined the components (a factor analysis can be undertaken by plotting the final RSS at the end of each component MCR and performing a Cattell scree test to estimate their number) their location in the data can be determined by using them as the start of a further round MCR-

ALS. Here, the original ToF-SIMS data were tessellated into 16 blocks, each 230×230 pixels, and a single iteration of MCR-ALS undertaken to indicate the spatial distribution of the components in the full-resolution image data.

B. Primary least squares regression analysis

Correlations between ToF-SIMS spectra and protein adsorption were assessed using PLS regression analysis as previously described.[20] In total, 706 negative ions were selected from a group of 70 polymers from the array. Ion peak intensities were dead time corrected,[39] normalized to the respective total secondary ion counts to remove the influence of primary ion beam fluctuation, mean-centered, and square root mean scaled prior to analysis.[4,38,40] PLS analysis was carried out using PLS Toolbox 5.2 software (Eigenvector). The dataset was randomly split into a training group, containing 75% of the samples, and a test set, containing the remaining 25% of samples. A "leave one out" cross validation method was used in the PLS regression analysis of the training set. The PLS model for the training set was validated by applying it to the test set. A sparse subset[41] of SIMS ions was produced for the final PLS model by initially producing a PLS model with all 706 peaks. Any peaks with a regression coefficient below 10% of the maximum regression coefficient calculated were removed and the PLS model was recalculated. This was repeated four times, after which 62 peaks were selected and used to produce a final model.

IV. RESULTS AND DISCUSSION

A. Array formation

A polymer microarray containing 70 distinct homopolymer chemistries (Table SI-1) was produced as illustrated schematically in Fig. 1(a) to evaluate the potential application of large area MCR imaging analysis of ToF-SIMS hyperspectral data. The polymer library consisted of a range of mono- and multifunctional polyacrylates and poly(meth)acrylates with a diverse range of hydrocarbon pendant groups. An overall assessment of the array printing quality was established by ToF-SIMS analysis using a selection of secondary ion images, including the total secondary ion image [Fig. 1(b)]. The "patchwork" appearance of the total ion image with demarcation gridlines at $500 \, \mu m$ intervals is an artifact of the image stitching process. The high quality of the array printing was determined by both optical microscopy (data not shown) and ToF-SIMS imaging analysis, with spot areas generally exhibiting good circularity and chemical uniformity. All 70 distinct spot regions corresponding to 70 unique polymers were observed along with a spot (71) where the solvent (DMF) was printed without a monomer. The appearance of the spot within the secondary ion images is likely due to topographic changes due to DMF remodeling of the pHEMA substrate. A relatively minor amount of chemical leaching[29] can be seen emanating from two spots, 32 and 35, which illustrates a "halo" around the edge of the spots [Fig. 1(b)].
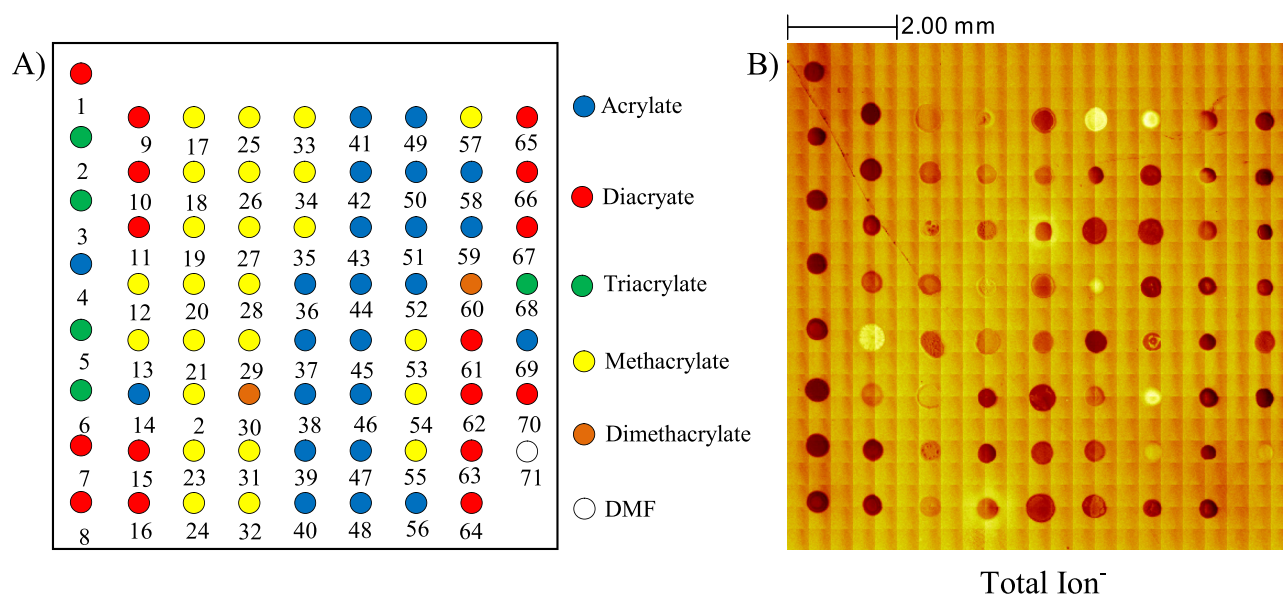
FIG. 1. (a) Schematic illustration of the microarray layout identifying polymer groupings of acrylate, diacrylate, triacrylate, methacrylate, and dimethacrylate spots and (b) total negative polarity secondary ion image of the 9.2×9.2 mm array area.

B. MCR imaging analysis

A variety of analytical methodologies have been applied in the deconvolution of the ToF-SIMS data acquired from polymer microarray systems, one of the most effective methods being MCR image analysis.[5,31,42] MCR image analysis was performed on the poly(meth)acrylate microarray dataset using a component number of 19, which was selected by the evaluation of the RSS, shown in Fig. SI-2. The scores images and loadings for all 20 components are shown in the supplementary material (Figs. SI-3 and SI-4, respectively) and illustrate the clear discrimination of a variety of chemistries specific to both polymer backbone and pendant groups as well as the pHEMA slide background (component 15). As a discussion of the key trends identified by the MCR image analysis, four components (11, 12, 17, and 19) have been presented in detail. The scores images and significantly loaded ions for these components are shown in Figs. 2(a) and 2(b), respectively. In the MCR analysis, each ion within the mass spectra is assigned a positive loading, the magnitude of which indicates how strongly correlated a particular ion is with a particular component. A secondary ion image for an example ion that was significantly loaded for each component is shown in Fig. 2(c).

The high intensity areas of component 17 correlated with the position of the monoacrylates within the array [shown in Fig. 1(a)]. The ion most highly loaded for this component was an acryloyl ion $C_3H_3O_2^-$, which likely originates from the polyacrylate backbone. Therefore, component 17 represents the monoacrylate chemistry within the polymer library. To validate the model the secondary ion images were assessed, and a pattern similar to that seen in component 17 [Figs. 2(a-i)] was observed in all cases, as is shown for the highly loaded $C_5H_5O^-$ secondary ion image shown in Fig. 2(c-i).

The distribution of component 19 [Fig. 2(a-ii)] correlated to the positions of the methacrylate polymer spots within the microarray. For this component a methacryloyl ion, $C_4H_5O_2^-$, was observed as the most highly loaded secondary ion [Fig. 2(b)]. It is therefore likely that this component is representative of the distribution of the methacrylate polymer backbone within the polymer microarray. As a validation of the model, the distribution of the significantly loaded ions for this component were investigated and found to closely match the MCR scores image for component 18, for example, the $C_4H_7O_4^-$ ion distribution shown in Fig. 2(c-ii). The backbone structure within the polymer library is a clear chemical discriminator and it is thus unsurprising that this aspect of the polymer chemistry is isolated as an MCR component. To complete the assessment of backbone structure, the position of the multiacrylate polymers is observed to correlate with MCR component 1 (Fig. SI-3).

In addition to identifying differences in the polymer backbone structure that are represented in all materials within the polymer library it was of interest to see whether the MCR analysis could identify components that represent side group chemical moieties. Component 13 did not correlate with any poly(meth)acrylate backbone structure, identifying a similarity in spots 26, 27, 28, 37, and 58 [Fig. 2(c-iii)]. The $C_6H_6O^-$ phenol ion was strongly loaded for this component. The spots most intense in the MCR scores image for component 13 were 27, 34, 37, and 58, and each of these polymers contains a phenol moiety within their pendant groups. The MCR analysis was therefore able to identify a specific chemical similarity between these materials. Two spots, 26 and 28, that do not have a phenol moiety within their pendant group showed a decreased intensity in component 13 compared with the four spots that contained a phenol group. The $C_6H_6O^-$ ion observed for these two materials is likely due to the fragmentation of the cyclic moiety to the resonance

A)   *MCR Scores Images*

| Component 17 | Component 19 | Component 13 | Component 11 |

B)   *Significantly Loaded Secondary ions*

	Component 17 (3.31 %)	Component 19 (3.83 %)	Component 13 (2.44 %)	Component 11 (1.04 %)
1	$C_3H_3O_2^-$	$C_4H_5O_2^-$	$C_6H_5O^-$	$C_3H_3O^-$
2	C_2HO^-	C_2H^-	C_4H^-	**$C_4H_3O_2^-$**
3	**$C_5H_5O^-$**	$C_3H_3O^-$	C_2H^-	$C_3H_5^-$
4	C_2H^-	CH^-	$C_4H_5O_2^-$	
5		**$C_4H_7O_4^-$**	**$C_6H_6O^-$**	
6			$C_3H_5O_2^-$	

C)   *ToF-SIMS Secondary Ion Images*

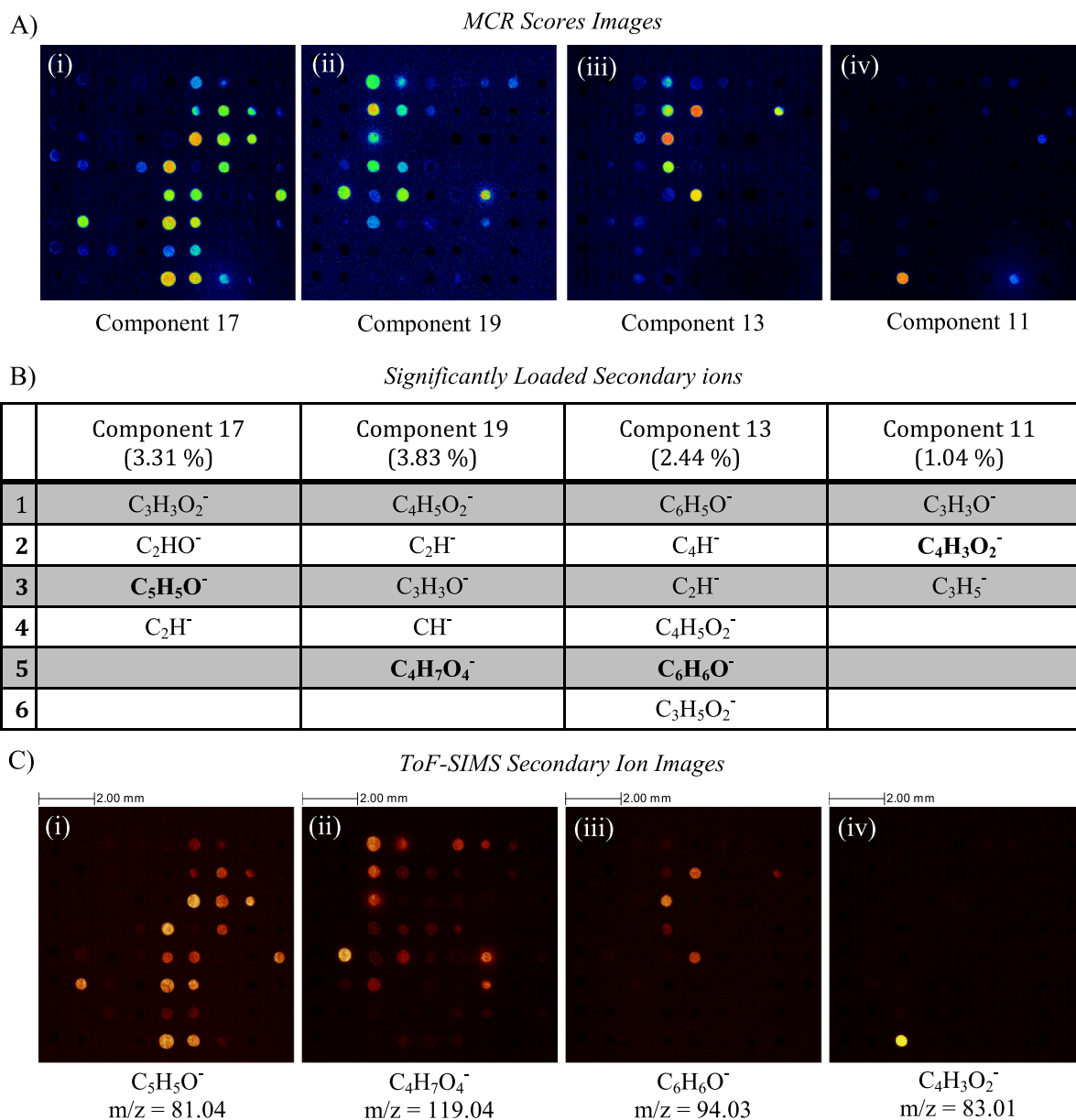| $C_5H_5O^-$ m/z = 81.04 | $C_4H_7O_4^-$ m/z = 119.04 | $C_6H_6O^-$ m/z = 94.03 | $C_4H_3O_2^-$ m/z = 83.01 |

FIG. 2. MCR image analysis showing four example scores component images for components: (i) 17, (ii) 19, (iii) 13, and (iv) 11, (b) their corresponding significantly highly loaded ions and (c) an example corresponding secondary ion image [highlighted in bold in (b)] for the (i) acrylate ($C_5H_5O^-$), (ii) methacrylate ($C_8H_7O^-$), (iii) phenyl side group ($C_6H_6O^-$), and (iv) acetoacetate side group ($C_5H_7O^-$) functional polymers.

stabilized phenol ion.[43] Other MCR components were found to correlate with chemical similarities with side-groups, for example, component 6 (Fig. SI-3) correlated with the polyethylene (PEG) side group located exclusively on polymers 32, 35, and 57.

The full extent to which MCR can deconvolute such complex datasets is epitomized in component 11 [Fig. 2(a-iv)], which differentiates a single polymer (spot 24) from the entire array. The highly loaded ion $C_4H_3O_2^-$ [Fig. 2(b)] likely originates from the acetoacetate group that is unique to spot 24 [Fig. 2(c-iv)]. Using this modified MCR analysis has allowed large area ToF-SIMS chemical imaging data to be successfully deconvoluted without compromising data or image resolution.

C. Protein adsorption

Further to the data exploration achieved using MCR image analysis, PLS regression allows the construction of structure-function models for the adsorption of albumin to the polymer microarray. The amount of fluorescently labeled protein strongly adsorbed on each polymer spot was quantified using a fluorescence scanner and the background fluorescence was removed. All of the data points were above the limit of detection. The data was then correlated with ToF-SIMS data extracted from individual spot ROI using PLS regression. The PLS data is summarized in Fig. 3. The 70 polymers were randomly split into a training (75%) and test set (25%). Root mean square error of cross validation
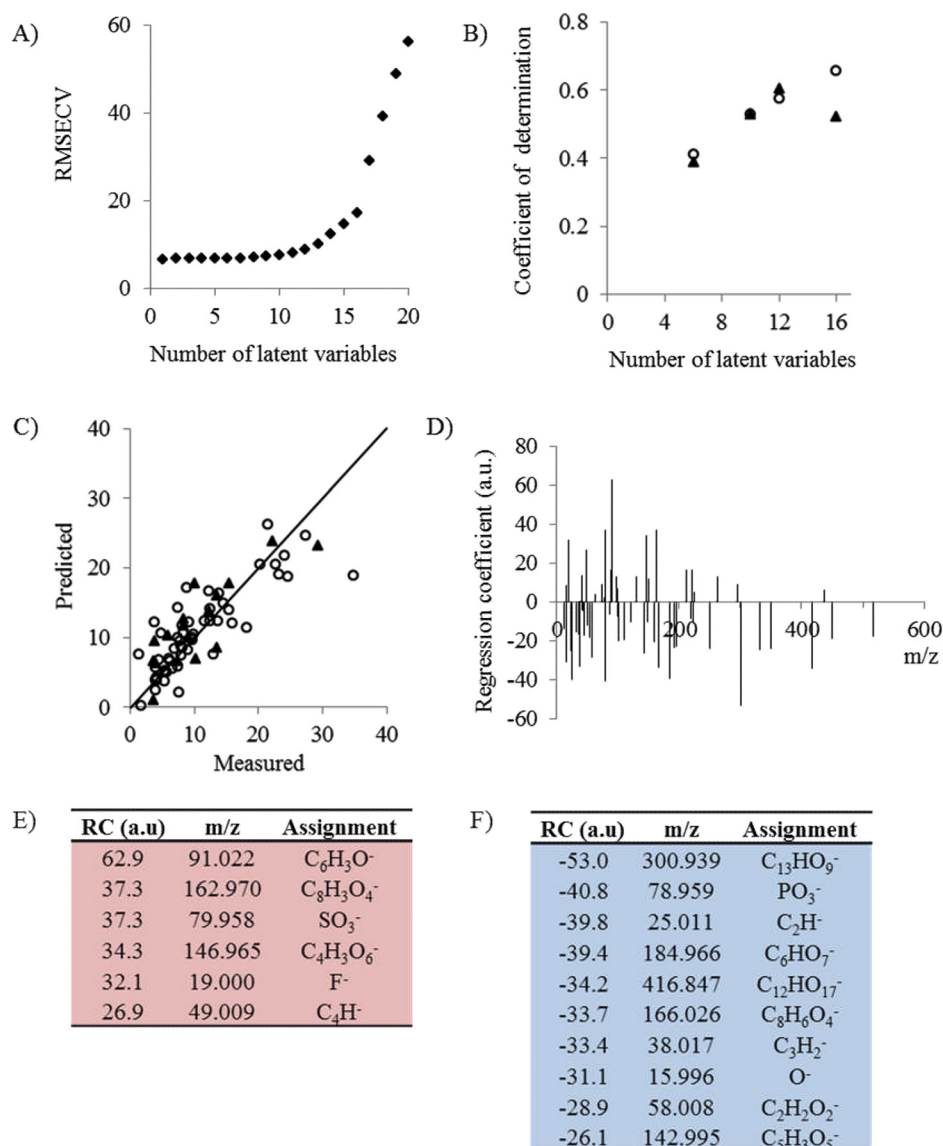
FIG. 3. Summary of PLS model used to predict albumin adsorption to the polymer library using conventional analysis. (a) The RMSECV determined for varied numbers of latent variables. (b) The coefficient of determination (R^2) determined for PLS models with varied numbers of latent variables for both the training (○) and test (▲) sets. (c) The measured vs predicted values determined by the PLS model for the training (○, $R^2 = 0.57$) and test (▲, $R^2 = 0.61$) sets. The y = x line has been drawn as a guide. (d) The regression vector for the PLS model. (e) and (f) Table of the secondary ions with the (e) highest and (f) lowest regression coefficients.

(RMSECV) curves were produced using the "leave one out" method as a guide to the number of latent variables to be used for the PLS model [Fig. 3(a)]. For further validation, multiple PLS models were produced from the training set using different numbers of latent variables, suggested by the RMSECV curves. These models were applied to the test and training sets and the R^2 value was calculated for the predicted values compared with the measured values [Fig. 3(b)]. The number of latent variables where the R^2 value for the measured versus predicted curves for the training and test sets was most similar was used to produce the final PLS models. A total of 12 latent variables were used to model protein adsorption. The R^2 value for the plot of the predicted versus the measured curve for protein adsorption was 0.57 for the training set and 0.61 [Fig. 3(c)]. The successful

prediction of protein adsorption to the polymer library, representative of a diverse chemical space, by assessing the surface chemistry of the materials represented in ToF-SIMS spectra is consistent with the established understanding that material chemistry is determinant on protein adsorption to a material.[44,45]

In the PLS model, each ion within the ToF-SIMS spectra is assigned a regression coefficient [Fig. 3(d)]. The ions that most strongly influence the model are identified by identifying ions with the largest positive [Fig. 3(e)] or negative [Fig. 3(f)] coefficient.

For the ions found to anticorrelate with protein adsorption [Fig. 3(e)], the $C_2H_2O_2^-$ ion also exhibited a high loading value for MCR component 6. This component correlated with the position of polymers containing a PEG side-group.
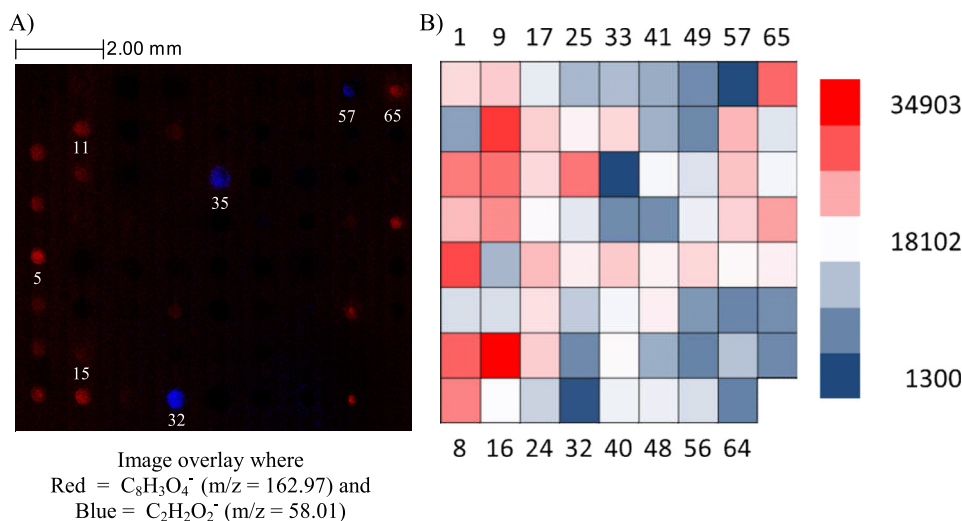
FIG. 4. (a) Secondary ion images illustrating the distribution of ions found to correlate ($C_8H_3O_4^-$) and anticorrelate ($C_2H_2O_2^-$) with protein adsorption and (b) the corresponding protein adsorption heat map determined from the fluorescence data for the array. A heat scale showing fluorescence in arbitrary units is shown to the right.

This observation is supported by the known antifouling properties of PEG.[46] It is likely that the $C_2H_2O_2^-$ originates from the PEG moiety and is specific to this group as opposed to short oligo(ethylene glycol) groups (n = 2, 3) for which a high intensity of this ion was not observed. When comparing the distribution of the $C_2H_2O_2^-$ ion to the fluorescence heat map (as shown in Fig. 4) it is clear that the three spots observed to show high intensities of PEG, 32, 35, and 57, correspond directly to the three most significantly correlating polymers with low protein adsorption regions.

The ions $C_8H_3O_4^-$ and $C_4H_3O_6^-$, which were assigned positive regression coefficients in the PLS model, were also associated with MCR component 1. This component correlated with multifunctional acrylate polymers (diacrylate, triacrylate, and dimethacrylate monomers). It is therefore likely that the ions $C_8H_3O_4^-$ and $C_4H_3O_6^-$ originated from the polymer backbone, suggesting that a surface enriched with acrylate backbone moieties promotes the adsorption of protein. The F^- and SO_3^- ions are due to the presence of contaminants on the polymer microarray. Comparison of the secondary ion images for the $C_8H_3O_4^-$ and $C_4H_3O_6^-$ with a heat map of protein adsorption (Fig. 4) revealed a correlation between the location of high protein adsorption and materials exhibiting high $C_8H_3O_4^-$ intensity. The four polymer spots exhibiting the highest protein adsorption, 5, 11, 15, and 65, can be clearly observed as some of the highest intensity regions for the $C_8H_3O_6^-$ ion. An acrylate backbone has previously been shown to increase the adsorption of protein to modulate stem cell attachment.[11]

The ability to validate observations made from such large datasets using PLS in Fig. 4 is a clear, rapid and novel approach utilizing the ToF-SIMS imaging data of the microarray. It is anticipated that this approach would prove successful for larger datasets including those without the well-ordered array format. Prior analysis using MCR imaging was also imperative to understanding the resultant PLS model.

V. SUMMARY AND CONCLUSIONS

In this study, a large area hyperspectral image ToF-SIMS dataset acquired from a poly(meth)acrylate polymer microarray has been successfully deconvoluted using a modified MCR analysis approach. This methodology can be used with a high specification desktop computer or with a high performance computer and allows the maintenance of the full resolution of the hyperspectral image datasets. This has led to the identification of chemical moieties associated with the polymer backbone and specific polymer side-groups. Using PLS regression analysis, the adsorption of protein to the polymer microarray was successfully predicted using only surface chemistry as represented by ToF-SIMS spectra with prior knowledge of the datasets from MCR analysis proving informative for interpretation. The novel application of the ToF-SIMS imaging datasets in the retrospective interpretation and validation of the PLS analysis has been demonstrated as an important tool for the further study of increasingly complex, multicomponent organic systems.

[1]J. Ibanez *et al.*, J. Appl. Phys. **108**, 103504 (2010).
[2]M. A. Robinson, D. J. Graham, and D. G. Castner, Anal. Chem. **84**, 4880 (2012).
[3]X. Y. Cui, D. A. Hutt, D. J. Scurr, and P. P. Conway, J. Electrochem. Soc. **158**, D172 (2011).
[4]D. J. Graham and D. G. Castner, Biointerphases **7**, 49 (2012).
[5]N. Scoutaris, A. L. Hook, P. R. Gellert, C. J. Roberts, M. R. Alexander, and D. J. Scurr, J. Mater. Sci. – Mater. Med. **23**, 385 (2012).
[6]W. Rao, D. J. Scurr, J. Burston, M. R. Alexander, and D. A. Barrett, Analyst **137**, 3946 (2012).
[7]A. L. Hook, D. G. Anderson, R. Langer, P. Williams, M. C. Davies, and M. R. Alexander, Biomaterials **31**, 187 (2010).

[8] A. J. Urquhart, M. Taylor, D. G. Anderson, R. Langer, M. C. Davies, and M. R. Alexander, Anal. Chem. **80**, 135 (2008).

[9] A. J. Urquhart, D. G. Anderson, M. Taylor, M. R. Alexander, R. Langer, and M. C. Davies, Adv. Mater. **19**, 2486 (2007).

[10] M. Taylor, A. J. Urquhart, D. G. Anderson, P. M. Williams, R. Langer, M. R. Alexander, and M. C. Davies, Macromol. Rapid Commun. **29**, 1298 (2008).

[11] Y. Mei *et al.*, Nat. Mater. **9**, 768 (2010).

[12] J. Yang *et al.*, Biomaterials **31**, 8827 (2010).

[13] R. Zhang *et al.*, Nat. Commun. **4**, 1335 (2013).

[14] A. L. Hook *et al.*, Nat. Biotechnol. **30**, 868 (2012).

[15] O. C. Sanni, D. G. Chang, R. Anderson, M. C. Langer, P. M. Davies, P. Williams, M. R. Alexander, and A. L. Hook, "Bacterial attachment to polymeric materials correlates with molecular flexibility and hydrophilicity," Adv. Healthcare Mater. (published online).

[16] S. Pernagallo, M. Wu, M. P. Gallagher, and M. Bradley, J. Mater. Chem. **21**, 96 (2011).

[17] R. S. Tare, F. Khan, G. Tourniaire, S. M. Morgan, M. Bradley, and R. O. C. Oreffo, Biomaterials **30**, 1045 (2009).

[18] E. Anglin *et al.*, Cytometry, Part A **77A**, 881 (2010).

[19] E. J. Anglin, C. Salisbury, S. Bailey, M. Hor, P. Macardle, M. Fenech, H. Thissen, and N. H. Voelcker, Lab Chip **10**, 3413 (2010).

[20] A. D. Celiz, H. C. Harrington, and A. L. Hook, Appl. Surf. Sci. **313**, 926 (2014).

[21] M. Taylor, A. J. Urquhart, M. Zelzer, M. C. Davies, and M. R. Alexander, Langmuir **23**, 6875 (2007).

[22] A. L. Hook, D. J. Scurr, D. G. Anderson, R. Langer, P. Williams, M. Davies, and M. Alexander, Surf. Interface Anal. **45**, 181 (2013).

[23] A. L. Hook, J. Yang, X. Y. Chen, C. J. Roberts, Y. Mei, D. G. Anderson, R. Langer, M. R. Alexander, and M. C. Davies, Soft Matter **7**, 7194 (2011).

[24] A. L. Hook, H. Thissen, and N. H. Voelcker, Langmuir **25**, 9173 (2009).

[25] C. A. Tweedie, D. G. Anderson, R. Langer, and K. J. Van Vliet, Adv. Mater. **17**, 2599 (2005).

[26] A. Chilkoti, A. E. Schmierer, V. Perezluna, and B. D. Ratner, Anal. Chem. **67**, 2883 (1995).

[27] M. C. Shen, M. S. Wagner, D. G. Castner, B. D. Ratner, and T. A. Horbett, Langmuir **19**, 1692 (2003).

[28] M. Taylor, A. J. Urquhart, D. G. Anderson, R. Langer, M. C. Davies, and M. R. Alexander, Surf. Interface Anal. **41**, 127 (2009).

[29] A. L. Hook, D. J. Scurr, J. C. Burley, R. Langer, D. G. Anderson, M. C. Davies, and M. R. Alexander, J. Mater. Chem. B **1**, 1035 (2013).

[30] A. M. Judd, D. J. Scurr, J. R. Heylings, K. W. Wan, and G. P. Moss, Pharm. Res. **30**, 1896 (2013).

[31] D. J. Scurr, A. L. Hook, J. Burley, P. M. Williams, D. G. Anderson, R. Langer, M. C. Davies, and M. R. Alexander, Surf. Interface Anal. **45**, 466 (2013).

[32] D. G. Anderson, S. Levenberg, and R. Langer, Nat. Biotechnol. **22**, 863 (2004).

[33] A. L. Hook *et al.*, J. Visualized Exp. **59**, e3636 (2012).

[34] See supplementary material at http://dx.doi.org/10.1116/1.4906484 for details of the monomers used in the study, a schematic depiction outlining the MCR-ALS process, the RSS plot used for determining the number of components in the MCR analysis, MCR scores images and significantly loaded secondary ions for all 20 MCR components.

[35] B. J. Tyler, Surf. Interface Anal. **46**, 581 (2014).

[36] K. M. Mullen, ALS: multivariate curve resolution alternating least squares (MCR-ALS), R package version 0.0.4, 2010, see http://CRAN.R-project.org/package=ALS.

[37] R Development Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2011, see http://www.R-project.org/.

[38] M. R. Keenan and P. G. Kotula, Surf. Interface Anal. **36**, 203 (2004).

[39] M. R. Keenan, V. S. Smentkowski, J. A. Ohlhausen, and P. G. Kotula, Surf. Interface Anal. **40**, 97 (2008).

[40] J. L. S. Lee, I. S. Gilmore, and M. P. Seah, Surf. Interface Anal. **40**, 1 (2008).

[41] T. Le, V. C. Epa, F. R. Burden, and D. A. Winkler, Chem. Rev. **112**, 2889 (2012).

[42] J. L. S. Lee, I. S. Gilmore, I. W. Fletcher, and M. P. Seah, Surf. Interface Anal. **41**, 653 (2009).

[43] W. J. Vanooij and R. H. G. Brinkhuis, Surf. Interface Anal. **11**, 430 (1988).

[44] K. L. Prime and G. M. Whitesides, Science **252**, 1164 (1991).

[45] Y. Arima and H. Iwata, Biomaterials **28**, 3074 (2007).

[46] P. Kingshott, S. McArthur, H. Thissen, D. G. Castner, and H. J. Griesser, Biomaterials **23**, 4775 (2002).