Anal. Chem. **2002,** *74,* 5711−5716

Principal Component Analysis of TOF-SIMS Images of Organic Monolayers

Mark C. Biesinger,*,† Pierre-Yves Paepegaey,† N. Stewart McIntyre,† Robert R. Harbottle,‡ and Nils O. Petersen‡

Surface Science Western and Department of Chemistry, The University of Western Ontario, London, Ontario, Canada N6A 5B7

Principal component analysis (PCA) is a statistical method used to find combinations of variables or factors that describe the most important trends in the data. PCA has been combined with time-of-flight secondary ion mass spectrometry (TOF-SIMS) data to extract new information and find relations between species contained in complex systems. Monolayers of dipalmitoylphosphatidylcholine alone and mixed with palmitoyloleoylphosphatidylglycerol prepared using the Langmuir−Blodgett technique are discussed. PCA software provides image scores and corresponding loadings for each significant principal component. Image plots of the scores show the spatial distribution and intensity of the species defined by the loading plots (mass spectral features). The intensity and resolution of the image scores can result in substantial improvement over that of the regular TOF-SIMS images especially when static conditions are used for small analysis areas. Also, some of the effects of topography and matrix in the images can be removed, allowing for a better presentation of chemical variations.

The use of principal component analysis (PCA) to analyze spectra of complex chemical systems is a relatively new investigative approach. At present, there have been a number of studies undertaken that combine time-of-flight secondary ion mass spectrometry (TOF-SIMS) spectral data and PCA. Lhoest et al.[1] used PCA to characterize TOF-SIMS fragmentation patterns of proteins according to protein and substrate type. Other examples include the use of PCA to study wash residues of food and beverage glass surfaces[2] and polystyrene spectra.[3] Examples of PCA combined with TOF-SIMS imaging data are rare at best. Kargacin and Kowalski[4] in 1986 laid out the foundations of PCA analysis on SIMS images and showed its ability to extract chemical information. However, due to the instrument quality and computing power available at that time, it is limited to a simple known system, low number of included data (12 mass-resolved images), and small image size (64 × 64 pixels). More recently, Lloyd et al.[5] combined PCA with TOF-SIMS imaging data to investigate surface-segregated additives in bulk polymers.

This paper will show that PCA, while an excellent tool to obtain trends in large amounts of data, also has the potential of increasing image contrast and resolution compared to normal TOF-SIMS images by combining the weighted intensities of related mass fragments into one image. This will be of particular use when small areas are imaged while ion beam dose rates are maintained below the static limit since any particular mass-resolved image is often restricted in its contrast and spatial resolution by low intensities (signal to background). PCA also has the potential to remove topographic and matrix effects from the data, allowing for a better understanding of true chemical variations at the surface of a sample.

Time-of-Flight Secondary Ion Mass Spectrometry. Due to its high lateral resolution, high mass resolution, wide mass range, and high sensitivity combined with low surface damage under static conditions, TOF-SIMS has acquired increasing importance in the past few years. This is especially true for the analysis of surfaces and films of organic materials. The TOF-SIMS techniques, described in detail elsewhere,[6,7] involve the use of a focused ion beam that bombards the surface of a sample causing the ejection of neutral and ionized species from the surface of the sample. Ionized species are collected and analyzed using a time-of-flight mass spectrometer. In the imaging mode, it is possible to raster the beam over the surface of the sample and collect a full mass spectrum for each point of analysis. Subsequently, images can be constructed for each secondary ion species corresponding to each mass detected. In complex systems, the number of possible species to be imaged may reach into the hundreds. It is often quite difficult to identify all of these species and to determine how they relate to each other. Also, higher mass fragments, though much more chemically significant, may have

* Corresponding author. E-mail: biesinger@julian.uwo.ca.
† Surface Science Western.
‡ Department of Chemistry.

(1) Lhoest, J. B.; Wagner, M. S.; Tidwell, C. D.; Castner, D. G.; *J. Biomed. Mater. Res.* **2001**, *57*, 432−440.
(2) Lang, F. *Time-of-Flight Secondary Ion Mass Spectrometry Analysis of Washing-Additives on Glass Surfaces*; SIMS XII; Elsevier: Amsterdam, 2000; pp 813−816.
(3) Vanden Eynde, X.; Bertrand, P. *Surf. Interface Anal.* **1997**, *25* (11), 878−888.
(4) Kargacin, M. E.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 2300−2306.
(5) Lloyd, K. G.; Stika, K. M.; Swartzfager, D. G. *Applications of Time-of-Flight SIMS in the Chemical Industry*; SIMS XI; Wiley: Chichester, 1998; pp 555−558.
(6) Wilson, R. G.; Stevie, F. A.; Magee, C. W. *Secondary Ion Mass Spectrometry*; Wiley: New York, 1989.
(7) Benninghoven, A.; Rüdenauer, F. G.; Werner, H. G. *Secondary Ion Mass Spectrometry: Basic Concepts, Instrumental Aspects, Application and Trends*; Wiley: New York, 1987.

10.1021/ac020311n CCC: $22.00
Published on Web 10/17/2002

a much weaker signal, especially if ion beam flux conditions must remain below the static limits for the sample.

Principal Component Analysis. A 256 × 256 pixel TOF-SIMS image data set obtained over a mass range of 1000 amu contains over 6.5×10^7 data points or variables. With such a large data set, extracting and analyzing the relevant data becomes a major issue. Fortunately, the variables are usually correlated such that the most important information is contained in a smaller number of components. Principal component analysis is employed to determine these components. PCA is a well-established technique and a full description of it can be found elsewhere.[4,8,9]

EXPERIMENTAL SECTION

Langmuir−Blodgett Films of DPPC and DPPC−POPG. Langmuir−Blodgett (LB) films of dipalmitoylphosphatidylcholine (DPPC) and a mixture of (3:1 by weight) DPPC and palmitoyloleoylphosphatidylglycerol (POPG) were prepared. LB films were deposited at pH 7 and 22 °C onto mica covered with a sputtered layer of gold. For the DPPC films, a surface pressure of 7 mN/m, a compression speed of 0.1 A^2 molecule^{-1} min^{-1}, and a transfer speed of 10 mm/min were used. For the DPPC−POPG mixture, a surface pressure of 20 mN/m, a compression speed of 0.5 A^2 molecule^{-1} min^{-1}, and a transfer speed of 10 mm/min were used. Monolayer films such as this have been shown to separate into two distinct regions, a liquid expanded (LE) phase, in which the hydrocarbon tails are fluid, and a liquid condensed (LC) phase, in which the tails are more rigid.[10,11] These features are readily discernible in the TOF-SIMS images.

TOF-SIMS Imaging Conditions. TOF-SIMS images were collected using a Cameca TOF-SIMS IV with a 25-keV gallium ion primary ion beam in "burst alignment" or medium-current mode. This mode uses a pulse width of 200 ns, spot size in the 250-nm range, and mass resolution of $M/\Delta M = 300$. All images obtained were 256 × 256 pixels.

PCA Software. PLS_Toolbox 2.1 from Eigenvector Research running on Matlab 6.0 was the software used for PCA analysis. For each set of data, as many significant mass peaks as possible were added to the peak list for analysis. For example, in the case of DPPC study, more than 60 are selected. Also included in the peak selection is the total remaining ion image (sum of ion intensity not selected as a specific peak) shown at mass zero in the loadings. This image is usually significant in that it contains a significant amount of topographic and matrix information. A conversion routine transforms the TOF-SIMS data (saved as a binary file) into a matrix usable by the PLS_Toolbox software. The data are then unfolded so that an image that was originally *I* by *J* pixels with *K* spectral channels is reshaped to form a 2-D array that is *I***J* by *K*. All data was "autoscaled" prior to PCA. In this procedure, the data are first mean-centered. This is done by subtracting the column mean from each column, thus forming a matrix where each column has a mean of zero. Each mean-centered variable is then divided by its standard deviation, which results in variables with unit variance. This procedure puts all variables on an equal basis in the analysis. Thus, the less intense but more chemically significant higher mass peaks receive the same level of consideration in the analysis as the intense, low-mass peaks. The principal component analysis model is then applied to the data.

After scores have been calculated for each pixel, the pixels are reorganized to the original image dimensions (*I* by *J*) and displayed as pseudocolor maps. This gives a graphic representation of the score value of each pixel as a function of position. Examination of the corresponding loadings for the each score image gives information about the various spectral variables (mass fragments) which are correlated with the score image.

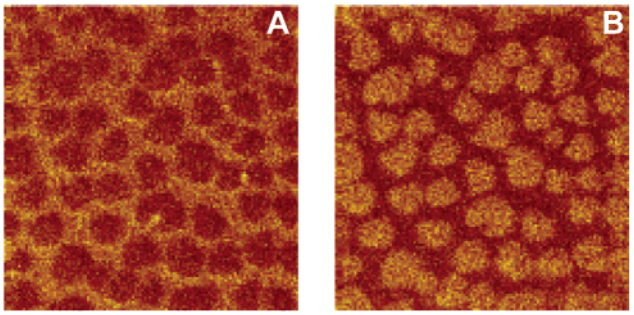

Figure 1. Total ion image of a DPPC sample before (A) and after (B) having passed the static limit. Images are 40 × 40 μm.

Table 1. Empirically Determined Values of the Static Limit for Langmuir−Blodgett Monolayers of DPPC and 3:1 DPPC−POPG

	ion dose (ions/cm^2)	
samples	200 × 200 μm	100 × 100 μm
DPPC	8×10^{12}	9×10^{12}
3:1 DPPC:POPG	8×10^{12}	$<9 \times 10^{12}$

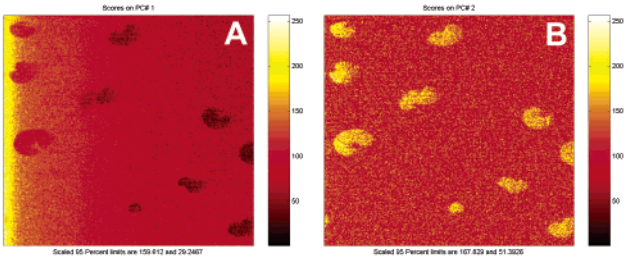

Figure 2. Image scores for PC 1 (A) and PC 2 (B) for DPPC. Images are 200 × 200 μm.

RESULTS AND DISCUSSION

Determination of the Static SIMS Limit. For the analysis of these Langmuir−Blodgett films, it was important that the static limit was not exceeded during the analysis. Initial studies have shown that large changes in the image contrasts can occur for these films as the static limit is approached and exceeded. Images of some mass fragments and even the total ion image can, after passing the static limit, appear as a negative image of the prelimit

(8) Mardia, K. V.; Kent, J. T.; Bibby, J. M. *Multivariate Analysis*; Academic Press: London, 1979; pp 213−254.
(9) Massart, D. L.; Vandeginste, B. G. M.; Buydens, L. M. C.; De Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J. *Handbook of Chemometrics and Qualimetrics*; Elsevier: Amsterdam, 1997; Part A, pp 519−556.
(10) Albrecht, O.; Gruler, H.; Sackmann, E. *J. Phys. (Paris)* **1978**, *39*, 301−313.
(11) McConnell, H. M.; Tamm, L. K.; Weis, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 3249

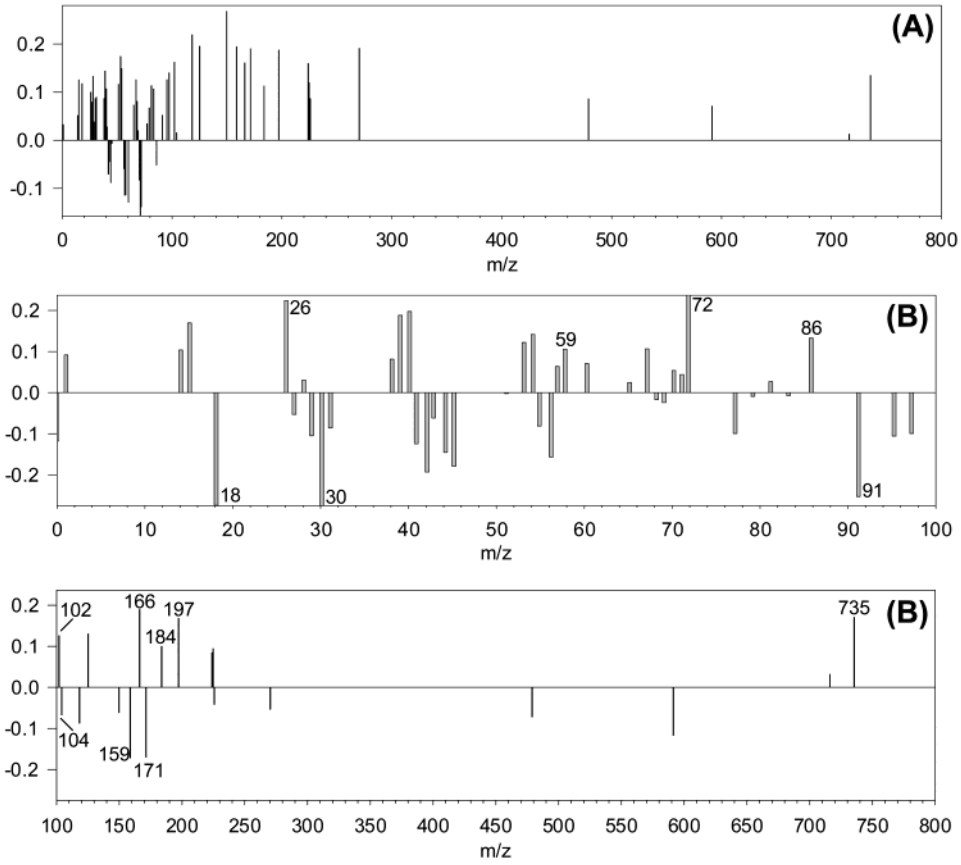

Figure 3. Loadings for PC 1 (A) and PC 2 (B) for DPPC.

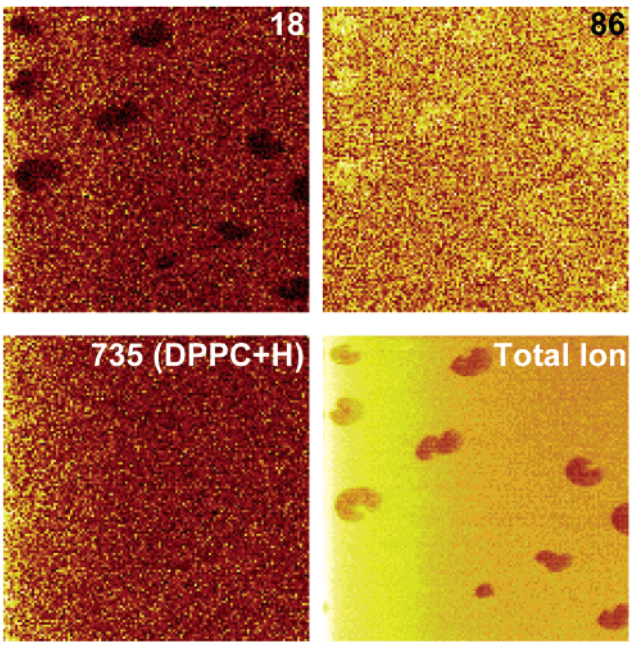

Figure 4. Selected TOF-SIMS images for the DPPC sample. Images are 200 × 200 μm.

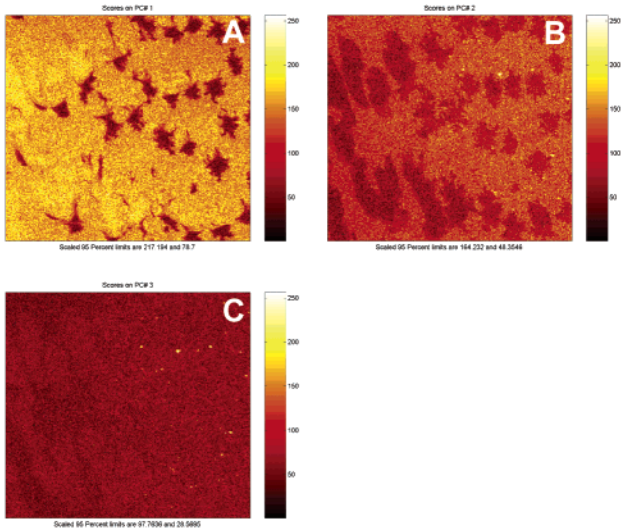

Figure 5. Image scores for PC 1 (A), PC 2 (B), and PC 3 (C) for DPPC−POPG. Images are 200 × 200 μm.

image (Figure 1). As a consequence, image contrast and resolution will be degraded and information pertaining to specific ion distributions will be corrupted. These changes can be abrupt, changing as soon as a specific amount of ion beam damage to the molecules in the sample has been reached. This effectively changes the matrix properties within the domains of the sample, causing ion yields to change dramatically. Effects like these need to be avoided; however, to maintain as much signal/noise as possible, one must attempt to approach as closely as possible the static limit.

Determination of the static limits was carried out as follows. Prepared films were imaged using either a 200 × 200 μm or a 100 × 100 μm area, a primary beam of known current, and a

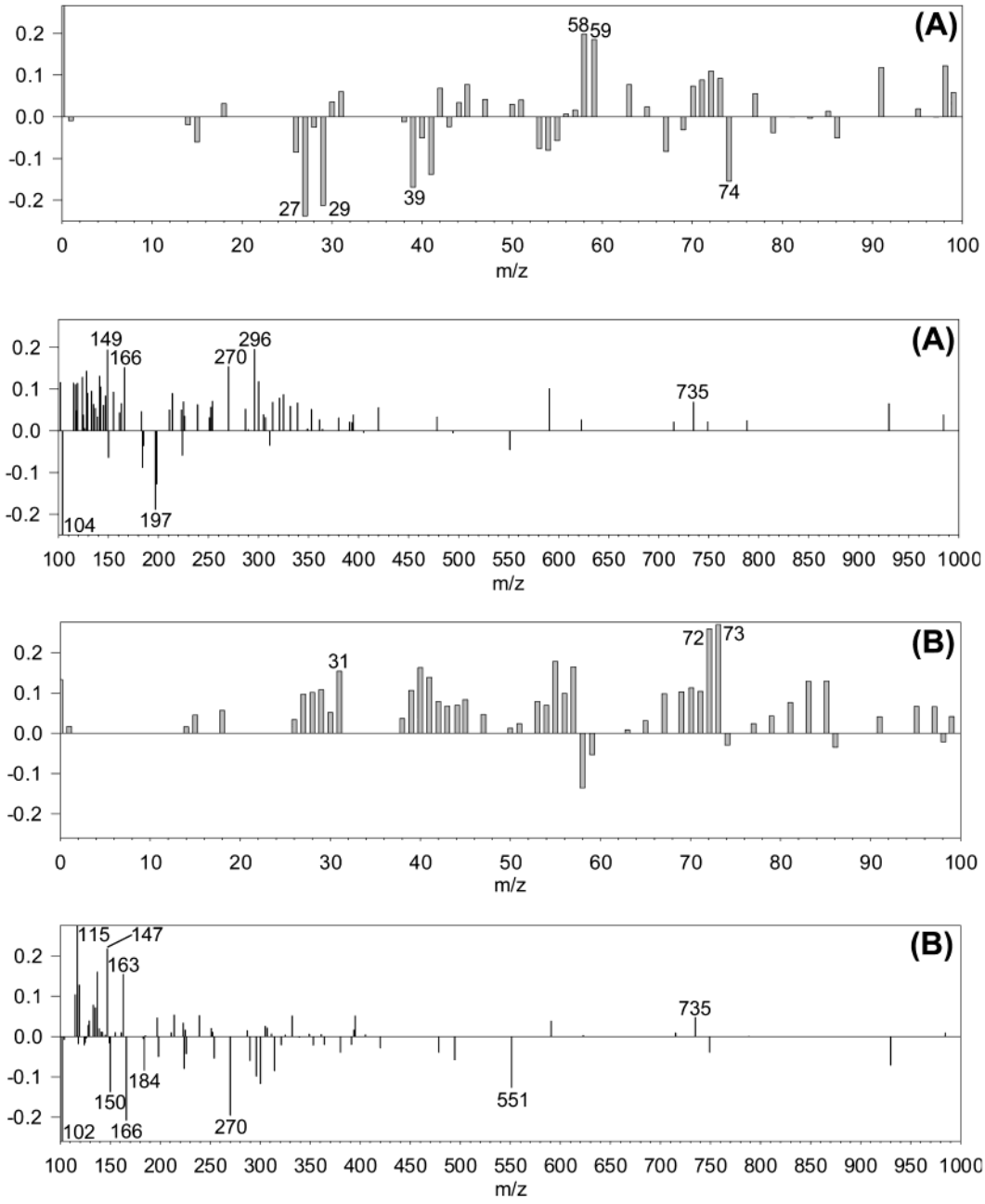

Figure 6. Loadings for PC 1 (A) and PC 2 (B) for DPPC−POPG.

known dwell time for each individual scan. Images were then reconstructed for each scan and the primary ion dose (ions/cm^2) calculated[12] for each scan. When any sign of a change in contrast occurred in the ion images, the static limit was considered to have been exceeded. Results presented in Table 1 show that static limits of between 10^{12} and 10^{13} ions/cm^2 are observed. This would correspond to a damage probability of between 1 and 10%, respectively.[12] During all subsequent analyses of these samples, a limit of $\sim 7 \times 10^{12}$ ions/cm^2 was not exceeded.

Langmuir−Blodgett Films of DPPC. Principal component analysis of 200 × 200 μm positive secondary ion TOF-SIMS images resulted in the identification of two principal components. Images scores are presented in Figure 2 with factor loadings presented in Figure 3. The factor loadings for PC 1 (Figure 3A) are quite strong over the entire mass range. This factor is thus ascribed to topographical effects that influence all or most of the mass peak intensities evenly. Topographical effects are exaggerated in this sample due to a slight tilt of the sample surface. This can be seen by a brightening of the left side in the image score for PC 1 (Figure 2A). As all mass fragments are affected by this effect, it accounts for a majority of the variance captured by PC 1.

Bourdos et al.[13] showed that TOF-SIMS image contrast variations in a single-component system are caused only by the

(12) Hagenhoff, B.; Rading, D. , Ion Beam Techniques: Surface Mass Spectrometry. In *Handbook of Surface and Interface Analysis: Methods for Problem Solving*; Rivière, J. C., Myhra, S., Eds.; Marcel Dekker Inc.: New York, 1998; pp 209−253.

(13) Bourdos, N.; Kollmer, F.; Benninghoven, A.; Sieber, M.; Galla, H. J. *Langmuir* **2000**, *16*, 1481−1484.

differing physical states of the lipids in the two phases. Generally, secondary ion intensities will decrease as the surface density of the molecules increase. Low-mass hydrocarbon fragments are much less affected by surface density whereas higher mass fragments can be greatly affected. However, most of the variations in ion yield seen for the area analyzed in general will also have been accounted for in PC 1 and thus will not be included in PC 2. Graham et al.[14] suggested that as the order in a self-assembled monolayer increases (more packed) one sees less low molecular weight species. Molecules laying down on the substrate are more susceptible to fragmentation. There may also be some suggestion of this in this work; however, it may depend on how loosely packed these molecules are in the LE phase. Graham's work was carried out using negative secondary ions where this work uses positive secondary ions. As well, the molecules used here have a somewhat more complicated structure than the dodecanethiol used by Graham. In particular, fragmentation from the headgroups of DPPC (and also POPG in the next section) may complicate this observation.

The image score for PC 2 (Figure 2B) shows quite clearly the LE and LC phases of DPPC within the sample. For PC 2, strong positive loading (Figure 3B) fragments (indicative of the LC phase, bright areas in Figure 2B) are found for masses such as 26 (CN), 72 ($C_4H_{10}N$), 86 ($C_5H_{12}N$), 102 ($C_5H_{12}NO$), 166 ($C_5H_{13}NPO_3$), 184 ($C_5H_{15}NPO_4$), 197 (Au), and 735 (DPPC + H). This indicates a general increase in the amount of DPPC in the LC phase and possibly for some ion yield variance not included in PC 1. Negative loadings include masses 18 (NH_4), 30 (CH_3NH), 91 (C_7H_7), 104 ($C_5H_{14}NO$), 159 (unassigned), and 171 (unassigned). These mass fragments are anticorrelated with the image score for PC 2.

In addition to PCA analysis removing a majority of topographic and ion yield or matrix effects, the image score for PC 2 has significantly better signal/noise than the individual mass-resolved TOF-SIMS images. The image score is effectively a combination of the weighted intensities of all mass fragments. This image improvement is critical since previous results have shown that obtaining high-quality images of chemically significant mass fragments can be difficult. Comparison of the image scores (Figure 2) to selected TOF-SIMS images (presented in Figure 4) dramatically shows this. This will be of even greater concern when chemically similar components are mixed. In normal TOF-SIMS imaging mode, only weaker intensity, higher mass fragments can be assigned with certainty to their species of origin.

Langmuir−Blodgett Films of DPPC−POPG. Films of (3:1) DPPC−POPG were analyzed in positive secondary ion mode using a 200 × 200 μm raster area. PCA analysis of the TOF-SIMS data resulted in three PCs (Figure 5). For the first PC, two different distributions can be distinguished. Bright but poorly defined structures are seen on the left side of the image while dark and sharply delineated structures are seen on the right side of the image. The loadings for PC 1 (Figure 6A) show mass spectral features that are correlated and anticorrelated with the image. Fragments with masses 58 (C_3H_8N), 59 (C_3H_9N), 149 (unassigned), 166 ($C_5H_{13}NPO_3$), 270 ($C_9H_{19}PO_7$), and 296 (unassigned) and the total remaining ion image (loading at mass zero) are all correlated with this image. Their corresponding TOF-SIMS images (see for example the image for mass 166 in Figure 7) will resemble PC 1 quite closely. A mass fragment such as mass 102 ($C_5H_{12}NO$), which has a weaker loading in PC 1, is less similar to the image score for PC 1. The TOF-SIMS image for mass 102 (Figure 7) shows the bright areas seen in PC 1 but does not show the corresponding dark areas. Mass peaks that have a negative correlation to PC 1 include 27 (C_2H_3), 29 (C_2H_5), 39 (C_3H_3), 74 (unassigned), 104 ($C_5H_{14}NO$), 197 (Au), and 198 (AuH). The TOF-SIMS images for these fragments (see, for example, the images for mass 27 and 104 in Figure 7) appear as a negative image of PC 1. Topographic effects do not appear to contribute greatly to the variance captured for PC 1. Thus, most of the effects seen can be attributed to ion yield effects and chemical differences that follow the structures seen in PC 1.

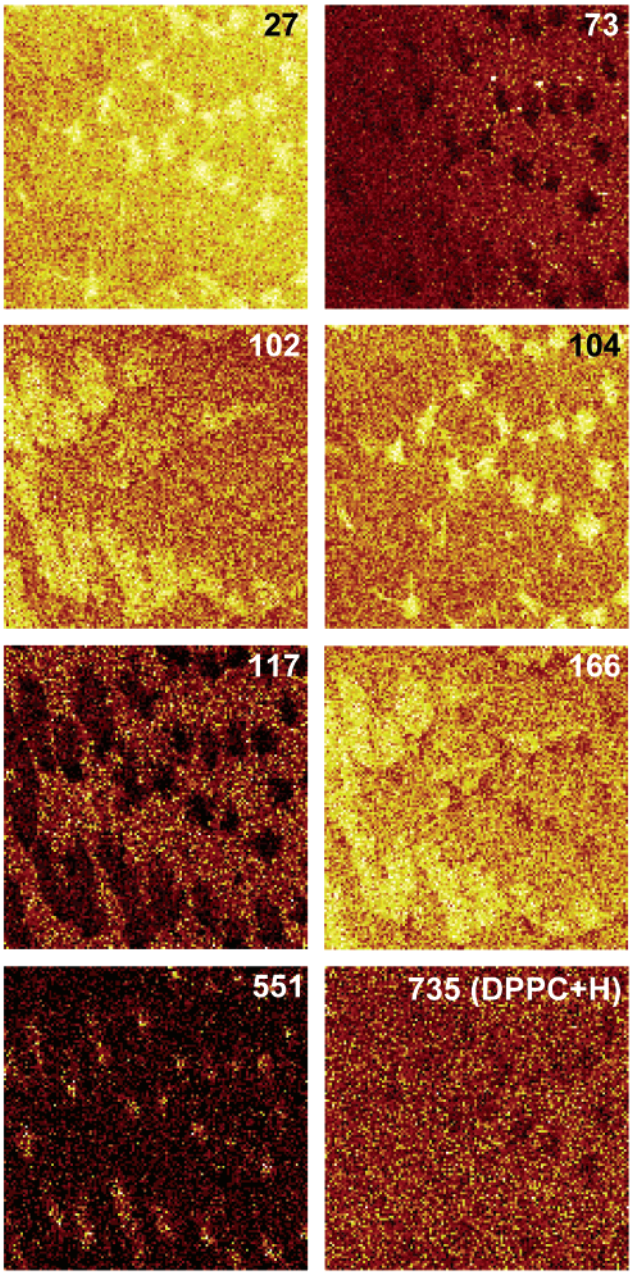

Figure 7. Selected TOF-SIMS images for the DPPC−POPG sample. Images are 200 × 200 μm.

(14) Graham, D. J.; Price, D. D.; Ratner, B. D. *Langmuir* **2002**, *18*, 1518−1527.

The second PC (loading presented in Figure 6B, image scores in Figure 5B), captures a relatively small amount of the variance in the data but gives some information about species that are present in the background, or LE phase, and species common to both bright and dark structures seen in PC 1. Species at mass 31 (P, CH_3O), 72 ($C_4H_{10}N$), 115 (unassigned), 117 (unassigned, see Figure 7), 119 (unassigned), and 163 (unassigned) are positively correlated with the image score for PC 2. The quasimolecular ion of DPPC + H (at mass 735), loads moderately strongly with both PC 1 and PC 2. A check of the corresponding TOF-SIMS image (Figure 7) shows that it has the dark structures seen in PC 1, but not the bright structures, making it somewhat similar to both PCs. Species at masses 102 ($C_5H_{12}NO$), 150 ($C_5H_{13}NPO_2$), 166 ($C_5H_{13}NPO_3$), 184 ($C_5H_{15}NPO_4$), 270 ($C_9H_{19}PO_7$), and 551 (DPPC + H − $C_5H_{15}NPO_4$) load negatively with PC 2. It should be noted that many of these mass fragments are also correlated with PC 1. However, PC 2 may be a better indication of strictly chemical phenomena in the sample than PC 1 is. This is due to the fact that most of the variance associated with ion yield effects has been accounted for in PC 1 and thus will not be seen in PC 2.

For PC 3, strong positive loadings for masses 73 ($(CH_3)_3Si$) and 147 ($(CH_3)_5Si_2O$) were found (loadings not shown). This indicates that there is a small amount of silicone (poly(dimethylsiloxane)) contamination on the surface of the sample. Some of this was also detected in PC 2, which also has strong positive loadings for these masses. The images for PC 2 and PC 3 (Figure 5C) both show a number of very small but intense areas (bright specks in PC 3) that likely correspond to silicone. The pattern for PC 3 also indicates that the larger, less sharply defined regions in the left of the image correspond to slightly higher concentrations of the silicone contamination.

PCA, as employed to this system, can be used as a tool to quickly associate structural features on the surface with specific mass fragments of importance. Using the principal component loadings in conjunction with TOF-SIMS images allows for an objective assessment of the information contained in the enormously large set of data. Some removal of ion yield effects was achieved and a trace contaminant, poly(dimethylsiloxane), was readily identified. Most importantly, the contrast and resolution of the image scores are a significant improvement relative to the mass-resolved TOF-SIMS images.

CONCLUSIONS

PCA is an excellent tool for deconvoluting complex sets of image data and is a good starting point to obtain information on trends within large data set. The contrast and resolution of the image scores are substantially improved over that of regular TOF-SIMS images when static conditions are used. Topographical and ion yield (matrix) effects can be removed effectively, allowing the analyst to concentrate on chemical variations across the surface within the system under study.

Received for review May 8, 2002. Accepted September 11, 2002.

AC020311N